

STIC-Biotech/ChemLib

121194

From: Rao, Manjunath N.  
Sent: Wednesday, May 05, 2004 10:05 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search request for 10/088,676

From: Manjunath N. Rao  
Art Unit 1652, Room 3B81  
Mail Box in Room 3C70  
Phone: 272-0939

RECEIVED  
MAY - 5 2004  
STIC

Date: 5-5-04

Please search the following as soon as possible for application with serial number

**10/088,676**

1. SEQ ID NO: 1 against all commercial nucleic acid databases, issued patents/published applications database and pending application database. Please provide a print of all results
2. SEQ ID NO: 2, against all commercial amino acid databases, issued patents/published applications database and pending application database. Please provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao  
Art Unit 1652, Room 3B81,  
Mail Box in Room 3C70,  
Remsen Building, USPTO

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 5/14  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

400, Dulany St.  
Alexandria, VA.  
Phone: 571-272-0939

RECEIVED  
MAY - 5 2004  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

Key	Location/Qualifiers
CDS	26..1804
	/*tag= a
	/product= "heparanase"

XX	MO20010643-A2.
PN	04-JAN-2001.
PB	
XX	19-JUN-2000; 2000WO-IL000358.
PF	
XX	
PR	25-JUN-1999; 99US-0140801P.
PA	(INSI-) INSIGHT STRATEGY & MARKETING LTD.
PI	
PL	Pecker I, Michal I, Itzhaki H;
PT	WPI; 2001-137930/14.
DR	P-PSDB; AAY97632.
XX	
FT	New polynucleotides and polypeptides that are distantly homologous to
PT	heparanase, useful in wound healing, as well as in gene therapy protocols
PP	for angiogenesis, restenosis, atherosclerosis, or inflammation.
PS	
XX	Claim 3; Fig 1; 67pp; English.
CC	This sequence encodes a heparanase of the invention. The heparanase DNA
CC	and protein sequences are useful in wound healing, angiogenesis,
CC	restenosis, atherosclerosis, inflammation, pulmonary diseases,
CC	neurodegenerative diseases (such as scrapie, Alzheimer's disease, and
CC	Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
CC	sequence is particularly useful in gene therapy
SQ	
	Sequence 2060 BP; 520 A; 557 C; 493 G; 490 T; 0 U; 0 Other:
	Query Match, 99.8%; Score 1775.8; DB 4; Length 2060;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1777; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
QY	1 ATGAGGGTCTTTGGCCTTCCTCGAAGCAATGCCCTCGACGAATCCGCCCCCGCGCG
DB	26 ATGAGGGTCTTTGGCCTTCCTCGAAGCAATGCCCTCGACGAATCCGCCCCCGCGCG
QY	61 TGCCCTAGCCCGGCGGACTCTCTACTTGAGCTCTGTGTGCATCTCTCCCTTCTCCGAG
DB	86 TGCCCTAGCCCGGCGGACTCTCTACTTGAGCTCTGTGTGCATCTCTCCCTTCTCCGAG
QY	121 GCTGGAACAAGAAGACCCTTGCTGCTTAACAGAGCTGCAAGTTTGAAGAAAAGACCTG
DB	146 GCTGGAACAAGAAGACCCTTGCTGCTTAACAGAGCTGCAAGTTTGAAGAAAAGACCTG
QY	181 ATTCTACTTGATGTAGGACCAAGAACCACTGACGACATCATATGAAATTCTCTCT
DB	206 ATTCTACTTGATGTAGGACCAAGAACCACTGACGACATCATATGAAATTCTCTCT
QY	241 CTGCACTGGAATCCGTCATCATTTGATATGAGCTGCTCGATTTCCTAACTCCAAGCG
DB	266 CTGCACTGGAATCCGTCATCATTTGATATGAGCTGCTCGATTTCCTAACTCCAAGCG
QY	301 TTGCTAACCTGAGCCCGGGAGATTTCGCGCCCTTTCTGCGCTTCGCGGGCAAAAGACC
DB	326 TTGCTAACCTGAGCCCGGGAGATTTCGCGCCCTTTCTGCGCTTCGCGGGCAAAAGACC
QY	361 GACTTCCTGACATTCAGAACCTTAGAAGCCCGGCGAAAAGCCGCGGGCCCGCGCCG
DB	386 GACTTCCTGACATTCAGAACCTTAGAAGCCCGGCGAAAAGCCCGCGGGCCCGCGCCG
QY	421 GATTACTATCTGAAAACTATGAGGATGACATTTGTGAAGTATGTTGCTTAGATAAA
DB	446 GATTACTATCTGAAAACTATGAGGATGACATTTGTGAAGTATGTTGCTTAGATAAA
QY	481 CAGAAAAGCTGCAAGATTGCCAGCACCTGTATGTATGCTGTCCAAAGAGAGAG
DB	506 CAGAAAAGCTGCAAGATTGCCAGCACCTGTATGTATGCTGTCCAAAGAGAGAG
QY	541 GCAGCTCAATGCAATCTGTTCTTTTAAAGAGCAATTCCTCAATATTTCAGTAATTC
DB	566 GCAGCTCAATGCAATCTGTTCTTTTAAAGAGCAATTCCTCAATATTTCAGTAATTC

QY 601 ATATTACAGCCAGGCTCTCTAGACAACTTTATTACTTTGCTGATGCTCTGAC 660  
 DB 626 ATTTTAAACCCAGGCTCTCTAGACAACTTTATTACTTTGCTGATGCTCTGAC 685  
 QY 661 CTGATATTTGCTCTTAAATGCACTGCGTCTGATGCTCTGATGCTCTGATGCT 720  
 DB 686 CTGATATTTGCTCTTAAATGCACTGCGTCTGATGCTCTGATGCTCTGATGCT 745  
 QY 721 GCGCTGATGCTCTTAAATGCACTGCGTCTGATGCTCTGATGCTCTGATGCT 780  
 DB 746 GCGCTGATGCTCTTAAATGCACTGCGTCTGATGCTCTGATGCTCTGATGCT 805  
 QY 781 AATGACCAAAATTAATCTATGCACTGCGTCTGATGCTCTGATGCTCTGATGCT 840  
 DB 806 AATGACCAAAATTAATCTATGCACTGCGTCTGATGCTCTGATGCTCTGATGCT 865  
 QY 841 AAGGATTTACATCCAGCTGAAAGAGCTGTTGCAAGCCCATCCGATTTATTCAGAGCCAGC 900  
 DB 866 AAGGATTTACATCCAGCTGAAAGAGCTGTTGCAAGCCCATCCGATTTATTCAGAGCCAGC 925  
 QY 901 TTATATGCGCCCTTAATTTGGGCGCGGAGGAGAAATGCAATGCGCCCTCTGATGCTCTGATGCT 960  
 DB 926 TTATATGCGCCCTTAATTTGGGCGCGGAGGAGAAATGCAATGCGCCCTCTGATGCTCTGATGCT 985  
 QY 961 ATGAGAGTGGCAGAAAGTACAGTATGCACTGCTGCAACATTTGCTATGATGAGC 1020  
 DB 986 ATGAGAGTGGCAGAAAGTACAGTATGCACTGCTGCAACATTTGCTATGATGAGC 1045  
 QY 1021 CCGGTGCTCAAGGTATGATGCTCTGCAAACTCGGCTGTTGACACACTCTCTGACAG 1080  
 DB 1046 CCGGTGCTCAAGGTATGATGCTCTGCAAACTCGGCTGTTGACACACTCTCTGACAG 1105  
 QY 1081 ATTAGGAAATTCAGAAAGTGTATTAATCATATCATCCAGAAAGAAATTTGGCTTGA 1140  
 DB 1106 ATTAGGAAATTCAGAAAGTGTATTAATCATATCATCCAGAAAGAAATTTGGCTTGA 1165  
 QY 1141 GGTGTGTGACCACTGCTGAGGAGCAAACTATTCGATTTCTATGCTGACAG 1200  
 DB 1166 GGTGTGTGACCACTGCTGAGGAGCAAACTATTCGATTTCTATGCTGACAG 1225  
 QY 1201 TTCTATGCTGACCACTTATGAAATGCTGCGCAATCGGAGATGATGCTGATGAG 1260  
 DB 1226 TTCTATGCTGACCACTTATGAAATGCTGCGCAATCGGAGATGATGCTGATGAG 1285  
 QY 1261 CACTCATTTTGTGACATGATGATCAATCACTCTGCTGACAGAAATTTTACCATTAACA 1320  
 DB 1286 CACTCATTTTGTGACATGATGATCAATCACTCTGCTGACAGAAATTTTACCATTAACA 1345  
 QY 1321 GACTATGCTGCTCTCTCTCTCAAAAGGCTGATGCGCCCAAAATGCTTGGCTGTCAT 1380  
 DB 1346 GACTATGCTGCTCTCTCTCTCAAAAGGCTGATGCGCCCAAAATGCTTGGCTGTCAT 1405  
 QY 1381 GTGCTGAGGCTCAAGGAGGAGCCGCTGCGAGTATCCGGGCAAACTAAGAGTT 1440  
 DB 1406 GTGCTGAGGCTCAAGGAGGAGCCGCTGCGAGTATCCGGGCAAACTAAGAGTT 1465  
 QY 1441 TATGCTACTGCAAAACCAACAACAACAACAATGCTGCTGCTGCTGCTGCTGCT 1500  
 DB 1466 TATGCTACTGCAAAACCAACAACAACAACAACAATGCTGCTGCTGCTGCTGCTGCT 1525  
 QY 1501 ATCATCAATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 DB 1526 ATCATCAATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1585  
 QY 1561 CTGCTTACAGTACTGCTGCTGAGGCTTATGCGAGGAGGCTTAAATGCTCAAGTCAAGT 1620  
 DB 1586 CTGCTTACAGTACTGCTGCTGAGGCTTATGCGAGGAGGCTTAAATGCTCAAGTCAAGT 1645  
 QY 1621 CAACGATATGCGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB 1646 CAACGATATGCGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705

QY 1681 GCGCCCTTGGGCGCGCGGACAT ACATGCGCTTTATG 1740  
 DB 1706 CCGCCCTTGGGCGCGCGGACATGCTCATCCCTCAGTACATGCGCTTTTGTG 1765  
 QY 1741 GTCAAGATGCTCAATGCTTTGGCGCGCTACAGATTA 1779  
 DB 1766 GTCAAGATGCTCAATGCTTTGGCGCGCTACAGATTA 1804

RESULT 4  
 AAH22671  
 ID AAH22671 standard; DNA; 2636 BP.  
 XX  
 XX AAH22671;  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Heparanase-like protein Hpa2 splice variant #1 encoding DNA.  
 KW Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KW cytosolic; neuroprotective; cerebroprotective; immunosuppressive;  
 KW antidiabetic; antiarteriosclerotic; vulnery; de.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..600  
 FT CDS /\*tag= a  
 FT 601..2379 /\*tag= b  
 FT /transl\_except= (pos: 1309..1311, aa: Xaa)  
 FT /note= "Xaa = unknown"  
 FT 3'UTR 2377..2636  
 FT /\*tag= c

MO200146392-A2.  
 28-JUN-2001.  
 21-DEC-2000; 2000NO-GB004963.  
 22-DEC-1999; 99GB-00030392.  
 07-APR-2000; 2000GB-0008713.  
 (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PI McKenzie BA, Stamps AC, Terrett JA, Tyson KL;  
 DR WPI; 2001-418056/44.  
 DR P-PSDB; AAB85215.  
 PT Novel homologs of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 XX  
 XX  
 PS Claim 12; Fig 1; 97pp; English.

The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the



sequence encoding it. Heparanase-2 is a member of the endoglycosyltransferase family of glycosyltransferases and it degrades heparan sulphate proteoglycans and HSPGs (biglycan) macromolecules of cell surfaces, basement membranes and the extracellular matrix. HSPGs support the vascular endothelium and stabilise the structure of the capillary wall. Heparanases may be associated with neovascularisation and metastasis related to malignant tumour formation. Heparanase-2 polynucleotides and proteins are useful as vaccines for inducing an immunological response against autoimmune disorders, blood coagulation disorders, cancer, diabetes, ischaemia, sepsis, stroke, cardiovascular diseases, or thrombosis, as well as in diagnosing (the susceptibility of a subject to) these diseases.

Heparanase-2 fragments may be used as immunogens to produce antibodies immunospecific to the polypeptides, and to identify membrane bound soluble receptors, agonists or antagonists that compete with the binding of the polypeptide to the receptor. An antibody specific for heparanase-2 can be used in the diagnosis of the above diseases and in isolating or identifying clones expressing heparanase-2. The present sequence represents heparanase-2. Three regions of heparanase-2 with high immunogenicity (immunogenic epitopes) can be used to raise antibodies against heparanase-2.

Sequence 592 AA;

Query Match	100.0%;	Score 3088;	DB 4;	Length 592;
Best Local Similarity	100.0%;	Pred. No. 1.3e-304;		
Matches 592; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	I	MRVYCAPPEAMPSNSRPPACLAGALYATALLIHLTSSQAGDRRLVVD5AAGIKETL	60
Dp	1	MRVYCAPPEAMPSNSRPPACLAGALYATALLIHLTSSQAGDRRLVVD5AAGIKETL	60
Qy	61	ILLDVSTKNRVTYNEHFLSLQDPSIIHDGMLDPLSSKRLVTLARGLSPAFIRFGKRT	120
Dp	61	ILLDVSTKNRVTYNEHFLSLQDPSIIHDGMLDPLSSKRLVTLARGLSPAFIRFGKRT	120
Qy	121	DLQFQNLRRNPAPKSRGGPGDYLLKNYEDDIVRSVDALDKQKCKIAQHPDVMVLQREK	180
Dp	121	DLQFQNLRRNPAPKSRGGPGDYLLKNYEDDIVRSVDALDKQKCKIAQHPDVMVLQREK	180
Qy	181	AAQMHVLLKEQSNYSNLIILABSLDKLNPACSSGLILFALNALRRPNNSNMSSS	240
Dp	181	AAQMHVLLKEQSNYSNLIILABSLDKLNPACSSGLILFALNALRRPNNSNMSSS	240
Qy	241	ALSLTKYSASKKNISWELGEBNNRYTHGAVNQSQJGXDYIOLKSLQPIRYSRAS	300
Dp	241	ALSLTKYSASKKNISWELGEBNNRYTHGAVNQSQJGXDYIOLKSLQPIRYSRAS	300
Qy	301	LYGNBGRPRKNYIALLDGFMKVASGTVAVTMQHCYIDGRVVKVMDPLKTRLLDPLSDQ	360
Dp	301	LYGNBGRPRKNYIALLDGFMKVASGTVAVTMQHCYIDGRVVKVMDPLKTRLLDPLSDQ	360
Qy	361	IRKIKQKVVNYYTGKKIKMLEGVTTSAGGNNLSDSYAGFMVMTLGMLANOGIDVIR	420
Dp	361	IRKIKQKVVNYYTGKKIKMLEGVTTSAGGNNLSDSYAGFMVMTLGMLANOGIDVIR	420
Qy	421	HSFEDHGNLVDQNFNP.LPDYMLSLLYKRLGPKYLAHVAGLQRPKRPGRVIRDXLRI	480
Dp	421	HSFEDHGNLVDQNFNP.LPDYMLSLLYKRLGPKYLAHVAGLQRPKRPGRVIRDXLRI	480
Qy	481	YAHCTNNHNHNVYRGSTIFLFINLHRSRKKIKLAGTLRDLKHQYLOLPYQOEGIKSKSV	540
Dp	481	YAHCTNNHNHNVYRGSTIFLFINLHRSRKKIKLAGTLRDLKHQYLOLPYQOEGIKSKSV	540
Qy	541	QLNQOPLVMDDGTLPELKRPLRAGGTVIPVTVMGFFVNVNNAALCRYR	592
Dp	541	QLNQOPLVMDDGTLPELKRPLRAGGTVIPVTVMGFFVNVNNAALCRYR	592

20-APR-2001 (first entry)  
Human heparanase, hnhpl, protein sequence.  
Heparanase; hnhpl; wound healing; angiogenesis; restenosis; Scarpe;  
atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;  
neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;  
gene therapy; human.  
Homo sapiens.  
WO200100643-A2.  
04-JAN-2001.  
19-JUN-2000; 2000WO-IL000358.  
25-JUN-1999; 99US-0140801P.  
(INST-) INSIGHT STRATEGY & MARKETING LTD.  
Pecker I, Michal I, Itzhaki H;  
WPI; 2001-137930/14.  
N-PSDB; AAA91097.  
New polynucleotides and polypeptides that are distantly homologous to  
heparanase, useful in wound healing, as well as in gene therapy protocols  
for angiogenesis, restenosis, atherosclerosis, or inflammation.  
Claim 10; Fig 1; 67pp; English.  
This sequence represents a heparanase of the invention. The heparanase  
DNA and protein sequences are useful in wound healing, angiogenesis,  
restenosis, atherosclerosis, inflammation, pulmonary diseases,  
neurodegenerative diseases (such as Scarpe, Alzheimer's disease, and  
Creutzfeldt-Jakob disease) or viral infections. The heparanase coding  
sequence is particularly useful in gene therapy

Query Match	99.7%	Score 3078	DB 4	Length 592
Best Local Similarity	99.7%	Pred. No. 1.3e-303		
Matches 590; Conservative	1	Mismatches	1	Indels 0; Gaps 0;

QY	1	MRVTCAPPEAMPSSNSRPPACTLAFGLYTLALLHLHSLSSQGDRRPLFVDBAAGLXKXTL	60
Db	1	MRVTCAPPEAMPSSNSRPPACTLAFGLYTLALLHLHSLSSQGDRRPLFVDBAAGLXKXTL	60
QY	61	ILLDVSTNCPRTYVENENFLSLQDPSIITHDGMFLDPSKRLVTLTARGLSPAFLRGGKRT	120
Db	61	ILLDVSTNCPRTYVENENFLSLQDPSIITHDGMFLDPSKRLVTLTARGLSPAFLRGGKRT	120
QY	121	DFLOFONLRNPAKSGGPGDPYYLKNYEDDIYRSDVALDKQKCKIAGHPDVMVLQBEK	180
Db	121	DFLOFONLRNPAKSGGPGDPYYLKNYEDDIYRSDVALDKQKCKIAGHPDVMVLQBEK	180
QY	181	AAQOMLVTLKQOFNSNTYENLLTARSGLDKYNEFADCSGLHLIPALNALRRPNNSWNSSS	240
Db	181	AAQOMLVTLKQOFNSNTYENLLTARSGLDKYNEFADCSGLHLIPALNALRRPNNSWNSSS	240
QY	241	ALSLTKYSASKKXNISMELAGNEPNNYRTMGRVAVNGSOLGQDYIOLKSLDPIRIRYSBAS	300
Db	241	ALSLTKYSASKKXNISMELAGNEPNNYRTMGRVAVNGSOLGQDYIOLKSLDPIRIRYSBAS	300
QY	301	LYGPNIGRPKNVIALLDGFMKVAGSTVDAYTMQHCYIDGRVYKMDFLKTRLLDTSDDQ	360
Db	301	LYGPNIGRPKNVIALLDGFMKVAGSTVDAYTMQHCYIDGRVYKMDFLKTRLLDTSDDQ	360
QY	361	TRKIOKRVNTYTPPGKKIWLGEVYTTSSAGCTNNLSVYAAAGELMNTLQMLANOGIGDIVIR	420
Db	361	TRKIOKRVNTYTPPGKKIWLGEVYTTSSAGCTNNLSVYAAAGELMNTLQMLANOGIGDIVIR	420

QY 421 HSPFDHGYNLVDONENPLDPYWLSTLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480  
 DB 421 HSPFDHGYNLVDONENPLDPYWLSTLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480  
 QY 481 YAHCTNNHNNYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGLSKSV 540  
 DB 481 YAHCTNNHNNYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGLSKSV 540  
 QY 541 QLNQGPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 592  
 DB 541 QLNQGPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 592

## RESULT 3

AAU07424

ID AAU07424 standard; protein; 592 AA.

AC AAU07424;

DT 18-DEC-2001 (first entry)

DE Human heparanase-like protein splice variant #1.

KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
 KW antiproliferative; cardiant; vasotropic; cerebroprotective; nocotropic;  
 KW neuroprotective; antibacterial; virocidic; fungicide; ophthalmological;  
 KW extracellular matrix; ECM; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection;  
 KW nervous system disorder; Alzheimer's disease; ocular disorder; sunburn;  
 KW wound healing; food additive; heparanase.

OS Homo sapiens.

PN MO200179253-A1.

PD 25-OCT-2001.

PF 11-APR-2001; 2001WO-US011643.

PR 18-APR-2000; 2000US-0198123P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Fiscella M, Shi Y, Ebner R, Ruben SM;

DR WPI; 2001-611720/70.

DR N-PSDB; AAS13848.

PT New nucleic acids encoding extracellular matrix polypeptides, for  
 PT diagnosing, treating, preventing or ameliorating human disorders and  
 PT disease, such as, autoimmune, hyperproliferative or cardiovascular  
 PT disorders.

PS Disclosure; Page 14; 308pp; English.

CC The invention relates to novel isolated polynucleotides (I) encoding  
 CC extracellular matrix (ECM) polypeptides. (I) and a polypeptide encoded by  
 CC (I) are used to prevent, treat or ameliorate a medical condition in e.g.  
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility to  
 CC a pathological condition. The antibodies to the polypeptides can also be  
 CC used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. The present sequence  
 CC represents the amino acid sequence of human heparanase-like protein,  
 CC splice variant #1

SQ Sequence 592 AA;

Query Match 99.7%; Score 3078; DB 4; Length 592;

Best Local Similarity 99.7%; Pred. No. 1.3e-303;

Matches 590; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRYLCAPFAPMBSNSRPPACIAPGALYIALIHLSLSQAGDRPLPDRAGLKEKTL 60  
 DB 1 MRYLCAPFAPMBSNSRPPACIAPGALYIALIHLSLSQAGDRPLPDRAGLKEKTL 60  
 QY 61 ILIDVSTKNPVTNENFLSLQDPSTIHGMDLFLSSRLVTLARGSLPAFLRFQKRT 120  
 DB 61 ILIDVSTKNPVTNENFLSLQDPSTIHGMDLFLSSRLVTLARGSLPAFLRFQKRT 120  
 QY 121 DFLQFQNLNRPASRGSGPGDYLLKNYEDDIVASDVALDKQCKIAQHPVWLVLQREK 180  
 DB 121 DFLQFQNLNRPASRGSGPGDYLLKNYEDDIVASDVALDKQCKIAQHPVWLVLQREK 180  
 QY 181 AAGMHLVLKEQSNNTYSNLIILFASLDKLYNADSGHLIFALNALRPNNSNNS 240  
 DB 181 AAGMHLVLKEQSNNTYSNLIILFASLDKLYNADSGHLIFALNALRPNNSNNS 240  
 QY 241 ALSILTKASAKKNIISWELGNEPNNTYMGRAVNSQGLQKDYIQKSLQPIRYSRAS 300  
 DB 241 ALSILTKASAKKNIISWELGNEPNNTYMGRAVNSQGLQKDYIQKSLQPIRYSRAS 300  
 QY 301 LYGNIGRPKNVIALIDGMKVASGVDAVTVQHCYIDGRVYKVMDFLTRLLDLSQ 360  
 DB 301 LYGNIGRPKNVIALIDGMKVASGVDAVTVQHCYIDGRVYKVMDFLTRLLDLSQ 360  
 QY 361 IRKIQKVVNTYTPGKIMLEGVTTGAGTNNISDSYAGFLMLNTLGMLANGIDVIR 420  
 DB 361 IRKIQKVVNTYTPGKIMLEGVTTGAGTNNISDSYAGFLMLNTLGMLANGIDVIR 420  
 QY 421 HSPFDHGYNLVDONENPLDPYWLSTLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480  
 DB 421 HSPFDHGYNLVDONENPLDPYWLSTLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480  
 QY 481 YAHCTNNHNNYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGLSKSV 540  
 DB 481 YAHCTNNHNNYVRSITLFTIINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGLSKSV 540  
 QY 541 QLNQGPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 592  
 DB 541 QLNQGPLVMVDDGTLPELKRPRLRAGRTLVIPVTMGFFVYKVNALACRYR 592

## RESULT 4

AAB85215

ID AAB85215 standard; protein; 592 AA.

AC AAB85215;

DT 07-SEP-2001 (first entry)

DE Heparanase-like protein Hpa2 splice variant #1.

KW Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;

KW cytosolic; neuroprotective; cerebroprotective; immunosuppressive;

KW antiproliferative; nocotropic; antiinflammatory; antiarthritic; antiaesthetic;

KW antidiabetic; antiarteriosclerotic; vulnary.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 237

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:32:35 ; Search time 724 Seconds  
(without alignments)  
10438.557 Million cell updates/sec

Title: US-10-088-676-1  
Perfect score: 1779  
Sequence: 1 atgagggtgcttgcgcctt.....tggcctgcgcctacgataa 1779

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseq1990s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002s:\*  
7: Geneseq2003as:\*  
8: Geneseq2003bs:\*  
9: Geneseq2003cs:\*  
10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1779	100.0	1779	4	AAf86101 Human Hep
2	1775.8	99.8	1779	4	AAf13848 DNA encod
3	1775.8	99.8	2060	4	AAf91097 Human hep
4	1773.8	99.7	2636	4	AAh22671 Hepatarnas
5	1769	99.4	1847	6	AAh29202 Human hep
6	1469.8	82.6	2496	6	AAf170849 Human hep
7	1443.4	81.1	1898	4	AAa91098 Human hep
8	1441.4	81.0	1685	6	AAh29204 Human hep
9	1417.8	79.7	1891	6	AAh21527 Human pol
10	1417.8	79.7	2326	6	AAf170705 Human pre
11	1417.8	79.7	2462	4	AAh22672 Hepatarnas
12	1411	79.3	1674	6	AAh29203 Human hep
13	1289.8	72.5	3943	5	AAh28347 Nucleotid
14	1132.2	63.6	2369	6	AAf170845 Human hep
15	1095.4	61.6	1724	4	AAa91099 Human hep
16	1095.4	61.6	1930	4	AAf13843 DNA encod
17	1095.4	61.6	2300	4	AAh22673 Hepatarnas
18	1093.4	61.5	1511	6	AAh29205 Human hep
19	833.4	47.0	1656	6	AAf170848 Human hep
20	686.2	38.6	1408	6	AAf170850 Human hep
21	632.4	35.5	1114	4	AAf198878 Human exc
22	632.4	35.5	1114	4	AAf164065 Human bla
23	624.8	35.1	1192	4	AAh98806 Human EST

24	432.4	24.3	468	4	AAf67040 Human hep
25	426.4	24.0	607	6	AAf170846 Rat hepar
26	334.4	18.8	336	4	AAf13849 DNA encod
27	292.8	18.5	1605	6	ABf40748 Chicken h
28	282.6	15.9	377	6	AAh26025 Human ORF
29	282.6	15.9	2396	4	AAa91113 Mouse hep
30	282.6	15.9	2396	4	AAa91113 Mouse hep
31	278.8	15.7	1721	2	AAx35648 DNA encd
32	278.8	15.7	1721	3	AAx35648 DNA encd
33	278.8	15.7	1721	3	AAa75051 CDNA encd
34	278.8	15.7	1721	3	AAa75051 CDNA encd
35	278.8	15.7	1721	4	AAa91112 Human hep
36	278.8	15.7	1899	2	AAx35650 CDNA encd
37	278.8	15.7	1899	2	AAa75053 CDNA encd
38	277.2	15.6	1669	7	AAh22816 Human hep
39	277.2	15.6	1669	9	ADf16011 G-coupled
40	277.2	15.6	1713	2	AAx37259 Human hep
41	277.2	15.6	1724	4	AAh20940 Human hep
42	277.2	15.6	3726	2	AAx86671 CDNA encd
43	277.2	15.6	3726	9	ADf18951 Human dis
44	275.6	15.5	1584	6	ABf40753 Chicken s
45	275.6	15.5	1593	2	AAf11236 Human pre

## ALIGNMENTS

RESULT 1	AAf86101 standard; cDNA; 1779 BP.
ID	AAf86101
AC	AAf86101;
XX	
XX	
DT	20-JUN-2001 (first entry)
XX	
XX	
DE	Human Heparanase-2 cDNA.
XX	
KW	Heparanase 2; human; endoglucuronidase; heparan sulphate; metastasis;
KM	neocarigigenesis; vaccine; autoimmune disorder; blood coagulation; ss;
KW	cancer; diabetes; ischaemia; sepsis; stroke; cardiovascular; thrombosis.
XX	
OS	Homo sapiens.
XX	
FM	Key
FT	CDS
FT	1..1779
FT	/*tag= a
FT	/product= "Heparanase-2"
XX	
XX	WC000121814-A1.
XX	
PD	29-MAR-2001.
XX	
PF	11-SEP-2000; 2000MC-EP008837.
XX	
PR	23-SEP-1999; 99EP-00118805.
XX	
PR	07-JUL-2000; 2000EP-00114649.
XX	
PA	(MERE ) MERCK PATENT GMBH.
XX	
PI	Duecker K, Strenberg C;
XX	
XX	WPI, 2001-308089/32.
DR	P-PDB; AAB81062.
XX	
PT	New heparanase-2 polypeptide useful in diagnosing (the susceptibility of
PT	a subject to) and as vaccines against e.g. autoimmune disorders,
PT	cardiovascular disease, cancer, diabetes, ischemia, sepsis, stroke, or
PT	thrombosis.
XX	
XX	Claim 5; Page 39-41; 46pp; English.
XX	
CC	This invention relates to a human Heparanase 2 protein and the cDNA
CC	sequence encoding it. Heparanase-2 is a member of the endoglucuronidase

CC family 5 polypeptides and it degrades heparan sulphate proteoglycans  
CC HSPGs (ubiquitous macromolecules of cell surfaces, basement membranes and  
CC the extracellular matrix). HSPGs support the vascular endothelium and  
CC stabilise the structure of the capillary wall. Heparanases may be  
CC associated with neovasculogenesis and metastasis related to malignant  
CC tumour formation. Heparanase-2 polynucleotides and proteins are useful as  
CC vaccines for inducing an immunological response against autoimmune  
CC disorders, blood coagulation disorders, cancer, diabetes, ischaemia,  
CC sepsis, stroke, cardiovascular diseases, or thrombosis, as well as in  
CC diagnosing (the susceptibility of a subject to) these diseases.  
CC Heparanase-2 fragments may be used as immunogens to produce antibodies  
CC immunospecific to the polypeptides, and to identify membrane bound  
CC soluble receptors, agonists or antagonists that compete with the binding  
CC of the polypeptide to the receptors. An antibody specific for heparanase-  
CC 2 can be used in the diagnosis of the above diseases and in isolating or  
CC identifying clones expressing heparanase-2. The present sequence  
CC represents cDNA encoding heparanase-2

Sequence 1779 BP; 443 A; 477 C; 437 G; 422 T; 0 U; 0 Other;

Query Match	Score	DB 4	Length
100.0%	1779	1779	1779

Matches 1779; Conservative 0; Mismatches

Matches 1779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	ATGAGGGTGGTTGGTCCCTCCGTGAAGCAATGCCCTACGACATCCCGGCCCCCGGG	60
Db	ATGAGGGTGGTTGGTCCCTCCGTGAAGCAATGCCCTACGACATCCCGGCCCCCGGG	60
Qy	1 ATGAGGGTGGTTGGTCCCTCCGTGAAGCAATGCCCTACGACATCCCGGCCCCCGGG	60
Db	1 ATGAGGGTGGTTGGTCCCTCCGTGAAGCAATGCCCTACGACATCCCGGCCCCCGGG	60
Qy	61 TGCCTAAGCCCCGGGGGGCTCTACTTGGCTGTGTGCTCATCTCTCCCTTCTCCAG	120
Db	61 TGCCTAAGCCCCGGGGGGCTCTACTTGGCTGTGTGCTCATCTCTCCCTTCTCCAG	120
Qy	121 GCCTGGAGACAGGAGACCCCTTGCCTGTGAACAAGCTGAGGTTTGAAGAAAAACCTCG	180
Db	121 GCCTGGAGACAGGAGACCCCTTGCCTGTGAACAAGCTGAGGTTTGAAGAAAAACCTCG	180
Qy	181 ATTCTACTTGATGTGAGCAACAAGAACCCAGTCAGSACATCAATGAAATTCCTCTGT	240
Db	181 ATTCTACTTGATGTGAGCAACAAGAACCCAGTCAGSACATCAATGAAATTCCTCTGT	240
Qy	241 CTGCAAGTGGATTCCTGTCATATCATGATGCTGCTCGATTTCTTAAGTCMAAGCG	300
Db	241 CTGCAAGTGGATTCCTGTCATATCATGATGCTGCTCGATTTCTTAAGTCMAAGCG	300
Qy	301 TTGGTGAACCTGGCCCGGGGACCTTTGGCCGCGCTTTCTGCGCTTCGGGGGCAAAAGACC	360
Db	301 TTGGTGAACCTGGCCCGGGGACCTTTGGCCGCGCTTTCTGCGCTTCGGGGGCAAAAGACC	360
Qy	361 GACTTCCTGCAGTTCCAGAACCTTAGAGAACCCCGGCAAAAGCCGCGGGGCCCG	420
Db	361 GACTTCCTGCAGTTCCAGAACCTTAGAGAACCCCGGCAAAAGCCGCGGGGCCCG	420
Qy	421 GATTCACTATCTCAAAAACATAAGAGATGACATGTGTGCAAGTGATGTGGCTTAAGTAA	480
Db	421 GATTCACTATCTCAAAAACATAAGAGATGACATGTGTGCAAGTGATGTGGCTTAAGTAA	480
Qy	481 CAGAAAGGCTGCAGAGATTGCCAGCACCTGTATGTATGTGTGCTCCAAAGGAGAG	540
Db	481 CAGAAAGGCTGCAGAGATTGCCAGCACCTGTATGTATGTGTGCTCCAAAGGAGAG	540
Qy	541 GCAGGCTCAGATGATGTGCTCTTCTTAAGAGAGCAATCTCCAAATCTTAAGATATTC	600
Db	541 GCAGGCTCAGATGATGTGCTCTTCTTAAGAGAGCAATCTCCAAATCTTAAGATATTC	600
Qy	601 AATTTAAACAGCCAGGCTCTGTAGACAAACCTTTATTAACCTTTGTGTATGCTCTGACATCCAC	660
Db	601 AATTTAAACAGCCAGGCTCTGTAGACAAACCTTTATTAACCTTTGTGTATGCTCTGACATCCAC	660
Qy	661 CTGATATTTGCTCTAAATGCACTGGGTGTATATCCCAATACTCTGTGAGACATGTTAAGT	720
Db	661 CTGATATTTGCTCTAAATGCACTGGGTGTATATCCCAATACTCTGTGAGACATGTTAAGT	720
Qy	721 GCCCTAGTCTGTGAAGTACAGCCGACCAAAAGTACCAATTTTGTGGAACTGGGT	780
Db	721 GCCCTAGTCTGTGAAGTACAGCCGACCAAAAGTACCAATTTTGTGGAACTGGGT	780

Db	721	GCCCTGAGTCTGTGAAGTACAGCCGACGAAAGTACAACTTTCTTGGGAATCGGT	780
OY	781	AATAGGCCAAATACTATCGGACCATGCGCGGACAGTAAATGCGACGCACTTGGGA	840
Db	781	AATAGGCCAAATACTATCGGACCATGCGCGGACAGTAAATGCGACGCACTTGGGA	840
OY	841	AAGATTACATCCAGCTGGAAGAGCCTGTGGAGCCCATCCGGATTATTCAGAGCCAGC	900
Db	841	AAGATTACATCCAGCTGGAAGAGCCTGTGGAGCCCATCCGGATTATTCAGAGCCAGC	900
OY	901	TTAATAGCCCTTAATATTGGCGCGGCCGAGAGAAATGTCATCGCCCTCTAGATGATTC	960
Db	901	TTAATAGCCCTTAATATTGGCGCGGCCGAGAGAAATGTCATCGCCCTCTAGATGATTC	960
OY	961	ATGAAAGTGGCAGAGAAATACAGTATGATGATTCCTGGCACTTCTCTACATTAATGAGC	1020
Db	961	ATGAAAGTGGCAGAGAAATACAGTATGATGATTCCTGGCACTTCTCTACATTAATGAGC	1020
OY	1021	CGGGTGGTCAAGGTGATGAGACTTCTGTAAAATCGCTGTGTAACAACATCTCTGACACG	1080
Db	1021	CGGGTGGTCAAGGTGATGAGACTTCTGTAAAATCGCTGTGTAACAACATCTCTGACACG	1080
OY	1081	ATTAGAAAATTCAGAAAAGTGTTAATAACATACACTCAGGAAAGAAATTTGGCTTGA	1140
Db	1081	ATTAGAAAATTCAGAAAAGTGTTAATAACATACACTCAGGAAAGAAATTTGGCTTGA	1140
OY	1141	GGTGTGTGACCACTCACTGAGTGGAGGCAAAACATATTCATTCATTTCTATGCTGACAGA	1200
Db	1141	GGTGTGTGACCACTCACTGAGTGGAGGCAAAACATATTCATTCATTTCTATGCTGACAGA	1200
OY	1201	TTCTTATGTTGGAACACTTTAGSANTCTGGCCATCAGGGCATTTGATGTGTGATACGG	1260
Db	1201	TTCTTATGTTGGAACACTTTAGSANTCTGGCCATCAGGGCATTTGATGTGTGATACGG	1260
OY	1261	CACATCATTTTGTGACATGAGTAATAATCACTCGTGGACACAGAAATTTTAAACCATTTACA	1320
Db	1261	CACATCATTTTGTGACATGAGTAATAATCACTCGTGGACACAGAAATTTTAAACCATTTACA	1320
OY	1321	GACTACTGGGCTCTCTCTCTCTCTACAAAGCGCTGATCGGCGCCCAAAAGTCTTGGCTGTGAT	1380
Db	1321	GACTACTGGGCTCTCTCTCTCTCTACAAAGCGCTGATCGGCGCCCAAAAGTCTTGGCTGTGAT	1380
OY	1381	GTTGGCTGGGCTCCAGCGGAAAGCCAGCGCTGGCGGAGTATCCGGGACAAACTTAAGGATT	1440
Db	1381	GTTGGCTGGGCTCCAGCGGAAAGCCAGCGCTGGCGGAGTATCCGGGACAAACTTAAGGATT	1440
OY	1441	TATGCTCATCTGCAAAACACACACACACACACTAGCTGTGGGTGCATTTACACTTTT	1500
Db	1441	TATGCTCATCTGCAAAACACACACACACACACTAGCTGTGGGTGCATTTACACTTTT	1500
OY	1501	ATCATCAACTGCATGCATCAGAAAGAAATAATCAGCTGGCTGGGACCTTCAGAGACAG	1560
Db	1501	ATCATCAACTGCATGCATCAGAAAGAAATAATCAGCTGGCTGGGACCTTCAGAGACAG	1560
OY	1561	CTGGTTCAACGATACCTGTGCAAGCCTTAAGGGCAGAGAGGGCTTAAAGTCCAATCACTG	1620
Db	1561	CTGGTTCAACGATACCTGTGCAAGCCTTAAAGGGCAGAGAGGGCTTAAAGTCCAATCACTG	1620
OY	1621	CAACTGATGCGCAGCCCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATG	1680
Db	1621	CAACTGATGCGCAGCCCTTAAGTATGATGATGATGATGATGATGATGATGATGATG	1680
OY	1681	CGCCCCCTCGGGCGGGCGGACATGATGATGATGATGATGATGATGATGATGATGATG	1740
Db	1681	CGCCCCCTCGGGCGGGCGGACATGATGATGATGATGATGATGATGATGATGATGATG	1740
OY	1741	GTCAGGAATGTCATGCTTTGGCCTGCGCGCTACCGATTA	1779
Db	1741	GTCAGGAATGTCATGCTTTGGCCTGCGCGCTACCGATTA	1779

## RESULT 2

AA513848  
ID AA513848 standard; CDNA; 1779 BP.  
XX  
AC AA513848;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE DNA encoding heparanase-like protein splice variant #1.  
XX  
Human; immunosuppressive; antidiabetic; antineoplastic; cytostatic;  
XX  
neuroprotective; cardiac; vasotropic; cerebroprotective; nocotropic;  
XX  
neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
XX  
extracellular matrix; ECM; autoimmune disease; rheumatoid arthritis;  
XX  
hyperproliferative disorder; neoplasm; cardiovascular disorder;  
XX  
cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection;  
XX  
nervous system disorder; Alzheimer's disease; ocular disorder; sunburn;  
XX  
wound healing; food additive; heparanase; ss.  
XX  
OS Homo sapiens.  
XX  
Key Location/Qualifiers  
XX  
CDS 1. 1779  
XX  
FT /product= "Heparanase-like protein, splice variant"  
XX  
FT  
XX  
PN WO200179253-A1.  
XX  
PD 25-OCT-2001.  
XX  
XX 11-APR-2001; 2001WO-US011643.  
XX  
PR 18-APR-2000; 2000US-0198123P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Fiselletta M, Shi Y, Ebner R, Ruben SM;  
XX  
DR MPI: 2001-611720/70.  
XX  
DR P-PSDB: AA007424.  
XX  
PT New nucleic acids encoding extracellular matrix polypeptides, for  
XX  
PT diagnosing, treating, preventing or ameliorating human disorders and  
XX  
PT disease, such as, autoimmune, hyperproliferative or cardiovascular  
XX  
PT disorders.  
XX  
PS Disclosure; Page 304-305; 308bp; English.  
XX  
XX The invention relates to novel isolated polynucleotides (I) encoding  
XX  
XX extracellular matrix (ECM) polypeptides. (I) and a polypeptide encoded by  
XX  
XX (I) are used to prevent, treat or ameliorate a medical condition in e.g.  
XX  
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They  
XX  
XX are also used in diagnosing a pathological condition or susceptibility to  
XX  
XX a pathological condition. The antibodies to the polypeptides can also be  
XX  
XX used in alleviating symptoms associated with the disorders and in  
XX  
XX diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
XX  
XX immunoassorbent assays (ELISA). Disorders which are diagnosed or treated  
XX  
XX include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
XX  
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
XX  
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
XX  
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
XX  
XX infections caused by bacteria, viruses and fungi and ocular disorders  
XX  
XX e.g. corneal infection. The polypeptides can also be used to aid wound  
XX  
XX healing and epithelial cell proliferation, to prevent skin aging due to  
XX  
XX sunburn, to maintain organs before transplantation, for supporting cell  
XX  
XX culture of primary tissues, to regenerate tissues and in chemotaxis. The  
XX  
XX polypeptides can also be used as a food additive or preservative to  
XX  
XX increase or decrease storage capabilities. The present sequence  
XX  
XX represents the coding sequence of heparanase-like protein splice variant  
XX  
XX #1  
XX  
SQ Sequence 1779 BP; 443 A; 477 C; 437 G; 422 T; 0 U; 0 Other;  
Query Match 99.8%; Score 1775.8; DB 4; Length 1779;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1777; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAGGCTGCTTGTGCTTCCCTTCCCTGAGGCAATGCCCTCCAGCAATCCCGCCCGCCGCG 60  
Db 1 ATGAGGCTGCTTGTGCTTCCCTTCCCTGAGGCAATGCCCTCCAGCAATCCCGCCCGCCGCG 60  
QY 61 TGCCTAGCCCGGGGCTCTACTTGGCTCTGTTGCTCCATCTCTCTCTCTCTCCAG 120  
Db 61 TGCCTAGCCCGGGGCTCTACTTGGCTCTGTTGCTCCATCTCTCTCTCTCTCCAG 120  
QY 121 GCTGAGACAGAGACCTTCTGCTGTAACAAGCTGACGTTTGAAGAAAAAGCCCTG 180  
Db 121 GCTGAGACAGAGACCTTCTGCTGTAACAAGCTGACGTTTGAAGAAAAAGCCCTG 180  
QY 181 ATTCTAGTATGATGAGACCAAGAACCCAGTCAGACAGTCATAGAACTTCTCTCT 240  
Db 181 ATTCTAGTATGATGAGACCAAGAACCCAGTCAGACAGTCATAGAACTTCTCTCT 240  
QY 241 CTGACGCTGATCCGATCATTCATTCATGATGCTGCTGATTTCTTAAGTCCAGGCG 300  
Db 241 CTGACGCTGATCCGATCATTCATTCATGATGCTGCTGATTTCTTAAGTCCAGGCG 300  
QY 301 TTGGTACCTGGCCCGGGGACTTTGGCCGCTTTCTGCGCTTGCGGGGCAAAAGGAC 360  
Db 301 TTGGTACCTGGCCCGGGGACTTTGGCCGCTTTCTGCGCTTGCGGGGCAAAAGGAC 360  
QY 361 GACTTCCTGATTCAGAACCTGAGAAACCGGCGAAAGCCGCGGGGCGCGCGCG 420  
Db 361 GACTTCCTGATTCAGAACCTGAGAAACCGGCGAAAGCCGCGGGGCGCGCGCG 420  
QY 421 GATTACTATCTCAAAACCTATGAGATGACATTTGTCGAGGATGCTTCTAGTAA 480  
Db 421 GATTACTATCTCAAAACCTATGAGATGACATTTGTCGAGGATGCTTCTAGTAA 480  
QY 481 CAGAAAGGCTGCAAGATGCTCCAGCAACCTGATGTTATGCTGCTGCTCCAAAGGAAG 540  
Db 481 CAGAAAGGCTGCAAGATGCTCCAGCAACCTGATGTTATGCTGCTGCTCCAAAGGAAG 540  
QY 541 GCAGCTGATGATCTGCTGTTCTTCTTAAGAGCAATTTCCAAATCTTACGATATC 600  
Db 541 GCAGCTGATGATCTGCTGTTCTTCTTAAGAGCAATTTCCAAATCTTACGATATC 600  
QY 601 ATATTACAGCCAGGCTCTAGACAACTTATATCTTGTGATGCTCTGAGCTCCAC 660  
Db 601 ATATTACAGCCAGGCTCTAGACAACTTATATCTTGTGATGCTCTGAGCTCCAC 660  
QY 661 CTGATATTTGCTTAATGACATGCGTGTATTCCTCAATPACTCTGGAACGTTCTAG 720  
Db 661 CTGATATTTGCTTAATGACATGCGTGTATTCCTCAATPACTCTGGAACGTTCTAG 720  
QY 721 GCCCTAGCTGTGTAAGTAGACGCGGCGCAAAATATACATTTCTGAGAACTGGGT 780  
Db 721 GCCCTAGCTGTGTAAGTAGACGCGGCGCAAAATATACATTTCTGAGAACTGGGT 780  
QY 781 AATGAGCAATAATATCTATGAGACATGCGCGGGGAGTAATGAGCAAGCTTGGGA 840  
Db 781 AATGAGCAATAATATCTATGAGACATGCGCGGGGAGTAATGAGCAAGCTTGGGA 840  
QY 841 AAGGATTAATCAAGTGAAGAGCCCTGTTGACGCCCATCCGATTTATCCAGGCAAG 900  
Db 841 AAGGATTAATCAAGTGAAGAGCCCTGTTGACGCCCATCCGATTTATCCAGGCAAG 900  
QY 901 TTATATGCTTATATTTGGGCGCGGAGAGAAATGATCATGCGCTCTAGATGATTC 960  
Db 901 TTATATGCTTATATTTGGGCGCGGAGAGAAATGATCATGCGCTCTAGATGATTC 960  
QY 961 ATGAGGTGACAGAGATCAAGTATGAGATGAGATTAAGTCAATGATGATGAGG 1020  
Db 961 ATGAGGTGACAGAGATCAAGTATGAGATGAGATTAAGTCAATGATGATGAGG 1020  
QY 1021 CGGCTGCTCAAGGTGATGATCTTCTGAAACTGCGCTGTAGACACATCTCTGACAG 1080  
Db 1021 CGGCTGCTCAAGGTGATGATCTTCTGAAACTGCGCTGTAGACACATCTCTGACAG 1080

```

Db      1021 CGGGTGTCAAGGATGACCTTCCTGAAAACCTGCTGTAGACACACTCTCTACACG 1080
Qy      1081 ATTAGGAAAAATTCAGAAAGTGTATATACATCTCCAGAGAAAGAAATTTGGCTTGA 1140
Db      1081 ATTAGGAAAAATTCAGAAAGTGTATATACATCTCCAGAGAAAGAAATTTGGCTTGA 1140
Qy      1141 GGTGTGTGACCACTGAGCTGAGGAGGACAAACATATATCGATTCATATGCTGAGAG 1200
Db      1141 GGTGTGTGACCACTGAGCTGAGGAGGACAAACATATATCGATTCATATGCTGAGAG 1200
Qy      1201 TTCTTAGTGTGAACACTTTAGGAATGCTGAGCAATGAGGCAATGATGCTGATACG 1260
Db      1201 TTCTTAGTGTGAACACTTTAGGAATGCTGAGCAATGAGGCAATGATGCTGATACG 1260
Qy      1261 CACTCATTTTGTGACCAATGATACATACCTGCTGAGACAGAAATTTTAACTTACCA 1320
Db      1261 CACTCATTTTGTGACCAATGATACATACCTGCTGAGACAGAAATTTTAACTTACCA 1320
Qy      1321 GACTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Db      1321 GACTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Qy      1381 GTGGCTGGGCTCCAGCGGAGAGCCAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1440
Db      1381 GTGGCTGGGCTCCAGCGGAGAGCCAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1440
Qy      1441 TATGCTCACTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1500
Db      1441 TATGCTCACTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1500
Qy      1501 ATCATCACTGATCGATCGATCAAGAAAGAAATCAAGCTGAGCTGAGCTGAGCTGAG 1560
Db      1501 ATCATCACTGATCGATCGATCAAGAAAGAAATCAAGCTGAGCTGAGCTGAGCTGAG 1560
Qy      1561 CTGGCTGACAGTACCTGCTGACAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Db      1561 CTGGCTGACAGTACCTGCTGACAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Qy      1621 CAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Db      1621 CAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Qy      1681 CGCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
Db      1681 CGCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
Qy      1741 GTCAAGATGTCATGCTTTGGCTGCGCTACCGATTA 1779
Db      1741 GTCAAGATGTCATGCTTTGGCTGCGCTACCGATTA 1779

RESULT 3
AAA91097
ID AAA91097 standard; DNA; 2060 BP.
AC AAA91097;
XX
XX 20-APR-2001 (first entry)
XX
XX Human heparanase, hnbp1, coding sequence.
XX
XX Heparanase; hnbp1; wound healing; angiogenesis; restenosis; Scurvy;
XX
XX atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;
XX
XX neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;
XX
XX gene therapy; human; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 26..1804
XX FT /*tag= a
XX FT /product= "heparanase"
XX

```

```

PN      W0200100643-A2.
XX
XX 04-JAN-2001.
XX
XX 19-JUN-2000; 2000MO-IL000358.
XX
XX 25-JUN-1999; 99US-0140801P.
XX
XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX
XX Pecker I, Michael I, Itzhaki H;
XX
XX WPI; 2001-137930/14.
XX
XX P-PSDB; AAY97632.
XX
XX New polynucleotides and polypeptides that are directly homologous to
XX
XX heparanase, useful in wound healing, as well as in gene therapy protocols
XX
XX for angiogenesis, restenosis, atherosclerosis, or inflammation.
XX
XX Claim 3; Fig 1; 67pp; English.
XX
XX This sequence encodes a heparanase of the invention. The heparanase DNA
XX
XX and protein sequences are useful in wound healing, angiogenesis,
XX
XX restenosis, atherosclerosis, inflammation, pulmonary disease, and
XX
XX neurodegenerative diseases (such as Scurvy, Alzheimer's disease, and
XX
XX Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
XX
XX sequence is particularly useful in gene therapy
XX
XX
XX Sequence 2060 BP; 520 A; 557 C; 493 G; 490 T; 0 U; 0 Other;
XX
XX
XX Query Match 99.8%; Score 1775.8; DB 4; Length 2060;
XX
XX Best Local Similarity 99.9%; Pred. No. 0;
XX
XX Matches 1777; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX
XX 1 ATGAGGGTGTGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
XX
XX 26 ATGAGGGTGTGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 85
XX
XX 61 TGGCTAGCCCGGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
XX
XX 86 TGGCTAGCCCGGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 145
XX
XX 121 GCTGAGACAGAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
XX
XX 146 GCTGAGACAGAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 205
XX
XX 181 ATTCTACTTGTATGAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAG 240
XX
XX 206 ATTCTACTTGTATGAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAG 265
XX
XX 241 CTGAGAGTGAATCGTCCATCATTCATGATGATGATGATGATGATGATGATGATGATGAT 300
XX
XX 266 CTGAGAGTGAATCGTCCATCATTCATGATGATGATGATGATGATGATGATGATGATGAT 325
XX
XX 301 TTGGTGAACCTGAGCCCGGGGACTTTCGCGCTTTCTGCGCTTTGGGGGCAAAAGAAC 360
XX
XX 326 TTGGTGAACCTGAGCCCGGGGACTTTCGCGCTTTCTGCGCTTTGGGGGCAAAAGAAC 385
XX
XX 361 GACTTCTGAGGTTGACAGACCTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAGAAC 420
XX
XX 386 GACTTCTGAGGTTGACAGACCTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAGAAC 445
XX
XX 421 GATTACTATCTCAAAACTATAGAGATGACATTTGTCAGAGTATGTTGCTTATGATTA 480
XX
XX 446 GATTACTATCTCAAAACTATAGAGATGACATTTGTCAGAGTATGTTGCTTATGATTA 505
XX
XX 481 CAGAAAGGCTGCAAGATGAGCCAGACCAAGCTGATGATGATGATGATGATGATGATGAT 540
XX
XX 506 CAGAAAGGCTGCAAGATGAGCCAGACCAAGCTGATGATGATGATGATGATGATGATGAT 565
XX
XX 541 GAGGCTCAGATCATCTGATTTCTTTAAAGAGCAATTTCTCAATACTTACAGTATCTC 600
XX
XX 566 GCAGCTCAGATCATCTGATTTCTTTAAAGAGCAATTTCTCAATACTTACAGTATCTC 625

```

QY 601 ATATTAAACAGCAGGTCTCTAGACAACTTATTAATTGTGATGCTCTGAGCTCCAC 660  
 DB 626 AATATTAAAGCCAGGCTCTAGACAACTTATTAATTGTGATGCTCTGAGCTCCAC 685  
 QY 661 CTGATATTTTGTCTTAATGACATGCGCTGCTAATCCCAATTAATCTCTGGAACGTTCTAGT 720  
 DB 686 CTGATATTTTGTCTTAATGACATGCGCTGCTAATCCCAATTAATCTCTGGAACGTTCTAGT 745  
 QY 721 GCGCTGAGTCTGTGAAGTACAGCGCCAGCAAAAATGACATTTCTTGGAACCTGGGT 780  
 DB 746 GCGCTGAGTCTGTGAAGTACAGCGCCAGCAAAAATGACATTTCTTGGAACCTGGGT 805  
 QY 781 AATGAGCCAAATTAATCTATGCGGACCATGATGCGCGGCGAGTAAATGCGACGCTTGGA 840  
 DB 806 AATGAGCCAAATTAATCTATGCGGACCATGATGCGCGGCGAGTAAATGCGACGCTTGGA 865  
 QY 841 AAGGATTTACATCAGGCTGAAGAGCCTGTTGACGCGCATCGGATTTAATCCAGAGCCAGC 900  
 DB 866 AAGGATTTACATCAGGCTGAAGAGCCTGTTGACGCGCATCGGATTTAATCCAGAGCCAGC 925  
 QY 901 TTATATGCGCCCTAATAATTGGGCGCGCGAGAGAGATGTCATCGCCCTCTAGATGATTC 960  
 DB 926 TTATATGCGCCCTAATAATTGGGCGCGCGAGAGAGATGTCATCGCCCTCTAGATGATTC 985  
 QY 961 ATGAGGTGCGAAGAGTATAGTAAATGAGTAACTGCTGCGACATGCTTAATGATGAGC 1020  
 DB 986 ATGAGGTGCGAAGAGTATAGTAAATGAGTAACTGCTGCGACATGCTTAATGATGAGC 1045  
 QY 1021 CCGGCTGCTCAAGGTGATGCACTTCTGAAAACTCGCCTGTTAGACACACTCTCTGACCG 1080  
 DB 1046 CCGGCTGCTCAAGGTGATGCACTTCTGAAAACTCGCCTGTTAGACACACTCTCTGACCG 1105  
 QY 1081 ATTAGGAAATTTCAAGAAAGTGTATTAATACATACATCTCCAGAGAAAGATTTGGCTTGA 1140  
 DB 1106 ATTAGGAAATTTCAAGAAAGTGTATTAATACATACATCTCCAGAGAAAGATTTGGCTTGA 1165  
 QY 1141 GGTGTGTGACACCTCAGCTGAGGCGCAAAACATCTATCCGATCTATGCTGACAGA 1200  
 DB 1166 GGTGTGTGACACCTCAGCTGAGGCGCAAAACATCTATCCGATCTATGCTGACAGA 1225  
 QY 1201 TTCTTATGTTGAACACTTTAGAGATGCTGGCCATTCAGGCGATTTGTCGTGATACG 1260  
 DB 1226 TTCTTATGTTGAACACTTTAGAGATGCTGGCCATTCAGGCGATTTGTCGTGATACG 1285  
 QY 1261 CACTCAATTTTTCAGCATGATCAATCACTCGTGGACAGAGATTTAACCACTTACCA 1320  
 DB 1286 CACTCAATTTTTCAGCATGATCAATCACTCGTGGACAGAGATTTAACCACTTACCA 1345  
 QY 1321 GACTACTGCTCTCTCTCTCTAACAAGGCTGATGCGCCCAAGTCTTGAGCTGTGACAT 1380  
 DB 1346 GACTACTGCTCTCTCTCTCTAACAAGGCTGATGCGCCCAAGTCTTGAGCTGTGACAT 1405  
 QY 1381 GTGCTGTGGCTCCAGCGGAGCCAGCGCTGCGGATGATCCGGGACAAACTTAAAGATTT 1440  
 DB 1406 GTGCTGTGGCTCCAGCGGAGCCAGCGCTGCGGATGATCCGGGACAAACTTAAAGATTT 1465  
 QY 1441 TATGCTCACTGCAAAACCAACCAACCACTAAGTGTGGGTCCATTAACCTTTTT 1500  
 DB 1466 TATGCTCACTGCAAAACCAACCAACCACTAAGTGTGGGTCCATTAACCTTTTT 1525  
 QY 1501 ATCATCAACTTTCATGATCAAGAAAGAAATCAAGTGTGGGTCCATTAACCTTTTT 1560  
 DB 1526 ATCATCAACTTTCATGATCAAGAAAGAAATCAAGTGTGGGTCCATTAACCTTTTT 1585  
 QY 1561 CTGCTTACACAGTACTGCTGACGCTTATGCGGACAGAGGCGCTAAAGTCCAAAGTCAAGTG 1620  
 DB 1586 CTGCTTACACAGTACTGCTGACGCTTATGCGGACAGAGGCGCTAAAGTCCAAAGTCAAGTG 1645  
 QY 1621 CAATGATGCGCGCCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB 1646 CAATGATGCGCGCCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1705

QY 1681 CGCCCCCTTGGCGCGCGCGGACATTTGATATCCCTCCAGTCAACATNGGCTTTATNG 1740  
 DB 1706 CGCCCCCTTGGCGCGCGCGGACATTTGATATCCCTCCAGTCAACATNGGCTTTATNG 1765  
 QY 1741 GTCAAGATGTCATGCTTTGGCTGCGCTGCTACCGATTA 1779  
 DB 1766 GTCAAGATGTCATGCTTTGGCTGCGCTGCTACCGATTA 1804  
 RESULT 4  
 ID AAH22671  
 XX AAH22671 standard; DNA; 2636 BP.  
 XX  
 AC AAH22671;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Heparanase-like protein Hpa2 splice variant #1 encoding DNA.  
 XX  
 KM Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KM cytosolic; neuroprotective; cerebroprotective; immunosuppressive;  
 KM antiproliferative; neurotropic; antiinflammatory; antiarthritic; antidiabetic;  
 KM antidiabetic; antiarteriosclerotic; vulnerrary; ds.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..600  
 FT 5'UTR  
 FT /\*tag= a  
 FT CDS 601..2379  
 FT /\*tag= b  
 FT /transl\_except= (pos: 1309..1311, aa: Xaa)  
 FT /note= "Xaa = unknown"  
 FT 3'UTR 2377..2636  
 FT /\*tag= c  
 PN WO200146392-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 21-DEC-2000; 2000MO-GB004963.  
 XX  
 PR 22-DEC-1999; 99GB-00030392.  
 PR 07-APR-2000; 2000GB-00008713.  
 XX  
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PI McKenzie EA, Stamps AC, Terrett JA, Tyson KL;  
 XX WPI; 2001-418056/44.  
 DR P-PSDB; AAB85215.  
 XX  
 PT Novel homologue of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 PS Claim 12; Fig 1; 97pp; English.  
 XX  
 CC The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the



CC nucleotide sequence of the largest splice variant of the heparanase-like  
 CC protein Hpa2 of the invention

XX Sequence 2636 BP; 707 A; 662 C; 673 G; 591 T; 0 U; 3 Other;

Query Match 99.7%; Score 1773.8; DB 4; Length 2636;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1775; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGGAGTGTGTGTGCTTCCCTGAAAGCATGCTCCAGCAACTCCGCCCCCGCG 60  
 Db 601 ATGAGGAGTGTGTGTGCTTCCCTGAAAGCATGCTCCAGCAACTCCGCCCCCGCG 660  
 QY 61 TGGCTTACCCCGGAGGCTCTTCACTTGGCTGTGTGCTCCATCTCTCCCTTCTCCAG 120  
 Db 661 TGGCTTACCCCGGAGGCTCTTCACTTGGCTGTGTGCTCCATCTCTCCCTTCTCCAG 720  
 QY 121 GCTGGAGACAGAGAGACCTTGGCTGTGAGACAGAGCTGAGGTTTGAAGAAAAGACCTG 180  
 Db 721 GCTGGAGACAGAGAGACCTTGGCTGTGAGACAGAGCTGAGGTTTGAAGAAAAGACCTG 780  
 QY 181 ATTCTAATTGATGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCTCT 240  
 Db 781 ATTCTAATTGATGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCTCT 840  
 QY 241 CTGACAGTGGATCCGTCCTCATCATTTGATGAGTGGCTGATTTCTTAAGCTCCAGACGC 300  
 Db 841 CTGACAGTGGATCCGTCCTCATCATTTGATGAGTGGCTGATTTCTTAAGCTCCAGACGC 900  
 QY 301 TTGGTGACCTGGCCCGGAGACTTTCGCGCTTTCGCGCTTTCGCGGAGCAAAAGAGAC 360  
 Db 901 TTGGTGACCTGGCCCGGAGACTTTCGCGCTTTCGCGCTTTCGCGGAGCAAAAGAGAC 960  
 QY 361 GACTTCTGAGATTCAGAACTGAGAAACCCGCGCAAAAGCCGCGGCGCCCGCG 420  
 Db 961 GACTTCTGAGATTCAGAACTGAGAAACCCGCGCAAAAGCCGCGGCGCCCGCG 1020  
 QY 421 GATTACTATCTCAAAAATCTATGAGATGACATTTGTCAGAGTGTTCCTTATGATAA 480  
 Db 1021 GATTACTATCTCAAAAATCTATGAGATGACATTTGTCAGAGTGTTCCTTATGATAA 1080  
 QY 481 CAGAAAGGCTGCAAGATTCGCCAGCACTGATGTTATGCTGTGCTCCAAAGGAGAG 540  
 Db 1081 CAGAAAGGCTGCAAGATTCGCCAGCACTGATGTTATGCTGTGCTCCAAAGGAGAG 1140  
 QY 541 GCAGCTCAGATGATCTGTTCTTCTAAAGAGCAATTCCTCAATCTTACAGTAATCTC 600  
 Db 1141 GCAGCTCAGATGATCTGTTCTTCTAAAGAGCAATTCCTCAATCTTACAGTAATCTC 1200  
 QY 601 ATATTAACAGCCAGTCTCTAGACAACTTTATTAATCTTGTGATGCTCTGACCTCAC 660  
 Db 1201 ATATTAACAGCCAGTCTCTAGACAACTTTATTAATCTTGTGATGCTCTGACCTCAC 1260  
 QY 661 CTATATTTGCTCTAAATGACAGTCGTCGTAATCCCAATATCTCCGGAACGTTTAT 720  
 Db 1261 CTATATTTGCTCTAAATGACAGTCGTCGTAATCCCAATATCTCCGGAACGTTTAT 1320  
 QY 721 GCCCTAGTCTGTTGAAGTACAGCGCCAGCAAAAGTACAACTTTCTTGGAACTGAGT 780  
 Db 1321 GCCCTAGTCTGTTGAAGTACAGCGCCAGCAAAAGTACAACTTTCTTGGAACTGAGT 1380  
 QY 781 AATGAGCAAAATTAATCTATGAGCAATGAGCGGCGGAGTAATAGCGACGATTTGGA 840  
 Db 1381 AATGAGCAAAATTAATCTATGAGCAATGAGCGGCGGAGTAATAGCGACGATTTGGA 1440  
 QY 841 AAGGATTACATCCAGCTGAAGACCTGTTGACAGCCCATCGGATTTATTCAGAGCCAGC 900  
 Db 1441 AAGGATTACATCCAGCTGAAGACCTGTTGACAGCCCATCGGATTTATTCAGAGCCAGC 1500  
 QY 901 TTATATGCGCCTAATTTGGGCGCGGAGAGAAATGATGATGCTTCTTATGATGATTC 960  
 Db 1501 TTATATGCGCCTAATTTGGGCGCGGAGAGAAATGATGATGCTTCTTATGATGATTC 1560

QY 961 ATGAGGAGTGGCAGAGATCAGTAGATGAGTTACTGTGCAACTTGTCTATGATGAC 1020  
 Db 1561 ATGAGGAGTGGCAGAGATCAGTAGATGAGTTACTGTGCAACTTGTCTATGATGATGAC 1620  
 QY 1021 CGGGTGGTCAAGGTGATGAGCTTCTGAAAACCTGGCTGTGTAGACACAGCTCTGACAG 1080  
 Db 1621 CGGGTGGTCAAGGTGATGAGCTTCTGAAAACCTGGCTGTGTGTAGACACAGCTCTGACAG 1680  
 QY 1081 ATTAGAAAATTGAGAAAGTGTAAATACATACATCTCCAGAAAAGAAATTGGCTTGA 1140  
 Db 1681 ATTAGAAAATTGAGAAAGTGTAAATACATACATCTCCAGAAAAGAAATTGGCTTGA 1740  
 QY 1141 GGTGTGTGACCACTGAGTGAAGGACAAATAATCTATCCATTCATGCTGAGAGA 1200  
 Db 1741 GGTGTGTGACCACTGAGTGAAGGACAAATAATCTATCCATTCATGCTGAGAGA 1800  
 QY 1201 TTCTTATGTTGAACAATTAGAAATGCTGGCAATCAGGGCAATTGATGCTGATACGG 1260  
 Db 1801 TTCTTATGTTGAACAATTAGAAATGCTGGCAATCAGGGCAATTGATGCTGATACGG 1860  
 QY 1261 CACTCATTTTTTGAACATGGAATCAATCACTCTGTGACCAATTTTAACCATTAACA 1320  
 Db 1861 CACTCATTTTTTGAACATGGAATCAATCACTCTGTGACCAATTTTAACCATTAACA 1920  
 QY 1321 GACTACTGGCTCTCTCTCTCTCAAGCGCTGATCGGCCCAAGTCTTGGCTGTGAT 1380  
 Db 1921 GACTACTGGCTCTCTCTCTCTCAAGCGCTGATCGGCCCAAGTCTTGGCTGTGAT 1980  
 QY 1381 GTGGCTGGGCTCCAGCGAAACCAAGCGCTGCGAGTATCCGGAGCAAACTTAAGATT 1440  
 Db 1981 GTGGCTGGGCTCCAGCGAAACCAAGCGCTGCGAGTATCCGGAGCAAACTTAAGATT 2040  
 QY 1441 TATGCTACAGCAAAACCAACCAACCAACCACTGCTGTGTGATCACTTTT 1500  
 Db 2041 TATGCTACAGCAAAACCAACCAACCAACCACTGCTGTGTGATCACTTTT 2100  
 QY 1501 ATCATCACTTGCATGATCAAGAAATCAAGCTGCTGGACTCTCAGAGCAAG 1560  
 Db 2101 ATCATCACTTGCATGATCAAGAAATCAAGCTGCTGGACTCTCAGAGCAAG 2160  
 QY 1561 CTGGTTACCAAGTCTCTGTCAGGCTTATGGGCAAGAGGCTTAAGTCAAGTCAAGT 1620  
 Db 2161 CTGGTTACCAAGTCTCTGTCAGGCTTATGGGCAAGAGGCTTAAGTCAAGTCAAGT 2220  
 QY 1621 CACTGATGCGCCAGCCCTTAGATGATGAGAGACCGGACCTCCAGATTTGAAGCC 1680  
 Db 2221 CACTGATGCGCCAGCCCTTAGATGATGAGAGACCGGACCTCCAGATTTGAAGCC 2280  
 QY 1681 CGCCCTCTTGGGCGGCGGAGCAATTTGATCTCTCCATCAACATGGGCTTTATGTG 1740  
 Db 2281 CGCCCTCTTGGGCGGCGGAGCAATTTGATCTCTCCATCAACATGGGCTTTATGTG 2340  
 QY 1741 GTCAAGATGTCATGCTTTGGCTGCGGCTACCGAATA 1779  
 Db 2341 GTCAAGATGTCATGCTTTGGCTGCGGCTACCGAATA 2379

# RESULT 5

AAD29202  
 ID AAD29202 standard; cDNA; 1847 BP.

XX AAD29202;

XX 07-MAY-2002 (first entry)

XX Human heparanase-2AB splice variant cDNA.

XX Human; heparanase-2AB; Hep-2; wound healing; angiogenesis; reestenosis;

XX atherosclerosis; neurodegenerative disease; inflammation; proteinase;

XX viral infection; autoimmune lesion; renal failure; pancreatic cancer;

XX dystrophic muscular disease; heart disease; gene therapy; enzyme; ss.

XX Homo sapiens.



XX Key Location/Qualifiers  
 PH CDS 85.1833  
 FT /tag= a  
 FT /product= "Human heparanase-2AB splice variant protein"  
 PN MO200204645-A2.  
 XX 17-JAN-2002.  
 PD 12-JUL-2001; 2001MO-EP008094.  
 PF 12-JUL-2000; 2000EP-00202442.  
 PR (VLA-) VLAMS INTERNUNIVERSITAIR INST BIOTECHNOG.  
 PA David G, Duerr J,  
 PI WPI; 2002-171719/22.  
 DR P-PSDB; AAE18326.  
 XX Heparanase-2 polypeptides and polynucleotides, useful for useful in wound  
 PT healing, angiogenesis, and for treating restenosis, atherosclerosis,  
 PT inflammation, neurodegenerative diseases, and viral infections.  
 XX Claim 2; Page 35-38; 54pp; English.  
 PS The invention relates to human heparanase-2 (Hep-2) polypeptides and  
 XX polynucleotides. Heparanase-2 protein is useful in wound healing  
 CC angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases,  
 CC inflammation and viral infections, as well as in neutralizing plasma  
 CC heparin as a potential replacement of procaine. Antiheparanase-2  
 CC antibodies may be used for immunodetection and diagnosis of  
 CC microvessel diseases, autoimmune lesions, renal failure in biopsy specimens,  
 CC plasma samples and body fluids. Molecules, which can agonise or  
 CC antagonise heparanase 2 catalytic activity may also be used as a  
 CC medicament. Polymorphisms in the polynucleotide sequence are useful in  
 CC the identification of individuals having a predisposition to acquire  
 CC diseases resulting from an increased or decreased expression of their  
 CC activity. Such molecules can be used to treat pancreatic cancer,  
 CC dystrophic muscular diseases and/or heart diseases. Polynucleotides of  
 CC the invention are used in gene therapy. The present sequence is human  
 CC heparanase-2AB splice variant cDNA  
 CC  
 XX Sequence 1847 BP; 465 A; 491 C; 456 G; 434 T; 0 U; 1 Other;  
 SQ  
 Query Match 99.4%; Score 1769; DB 6; Length 1847;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1772; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 361 GACTTCTGCAAGTTCAGAACTTGAGAACCCGGCCGAAAGCCGGGGGCGCCGGCCG 420  
 DB GACTTCTGCAAGTTCAGAACTTGAGAACCCGGCCGAAAGCCGGGGGCGCCGGCCG 474  
 QY 421 GATTACTATCTCAAAAACATAGAGATGATGATGTTGCGATTAGATGATGATGATG 480  
 DB GATTACTATCTCAAAAACATAGAGATGATGATGTTGCGATTAGATGATGATGATG 534  
 QY 481 CAGAAAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
 DB CAGAAAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 594  
 QY 541 GAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 DB GAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 654  
 QY 601 ATATTAAACAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 DB ATATTAAACAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 714  
 QY 661 CTGATATTGCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 DB CTGATATTGCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774  
 QY 721 GCGCTGAGTCTGTGAAGTACAGCCGAGCAAAAAGTACAACTTTCTTGGGAACTGGGT 780  
 DB GCGCTGAGTCTGTGAAGTACAGCCGAGCAAAAAGTACAACTTTCTTGGGAACTGGGT 834  
 QY 781 AATGAGCCAAATATCTATGCAATGATGATGATGATGATGATGATGATGATGATGATG 840  
 DB AATGAGCCAAATATCTATGCAATGATGATGATGATGATGATGATGATGATGATGATG 894  
 QY 841 AAGATTATCATCAGCTGAGAAAGCTGTGAGCCATCCGATTTATTCAGAGCCAGC 900  
 DB AAGATTATCATCAGCTGAGAAAGCTGTGAGCCATCCGATTTATTCAGAGCCAGC 954  
 QY 901 TTAATGAGCCCTAATATTTGGGCGGCGGAGGAAATGATGATGATGATGATGATGATG 960  
 DB TTAATGAGCCCTAATATTTGGGCGGCGGAGGAAATGATGATGATGATGATGATGATG 1014  
 QY 961 ATGAAGGTGAGGAGGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB ATGAAGGTGAGGAGGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGATG 1074  
 QY 1021 CGGTTGTCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 DB CGGTTGTCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1134  
 QY 1081 ATTAGGAATAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
 DB ATTAGGAATAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1194  
 QY 1141 GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
 DB GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254  
 QY 1195 GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
 DB GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1314  
 QY 1201 TTCTTAATGTTGAACATTTAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 1260  
 DB TTCTTAATGTTGAACATTTAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 1314  
 QY 1255 TTCTTAATGTTGAACATTTAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 DB TTCTTAATGTTGAACATTTAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 1374  
 QY 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 DB CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374  
 QY 1315 GACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
 DB GACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1434  
 QY 1375 GACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
 DB GACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1494  
 QY 1381 GTGGCTGGGCTCCAGCGGAAACCAAGGCTGAGGATGATGATGATGATGATGATGATGATG 1440  
 DB GTGGCTGGGCTCCAGCGGAAACCAAGGCTGAGGATGATGATGATGATGATGATGATGATG 1494



Fri May 14 13:01:53 2004

us-10-088-676-1.rng

**ପାଞ୍ଚ ୨**

Db	1189	GATCTTATGGTGAACACTTTAGGAATGCTGGCCAAATCAGGGCATATGTCGTATAC	1258																																																																									
Qy	1189	GATCTTATGGTGAACACTTTAGGAATGCTGGCCAAATCAGGGCATATGTCGTATAC	1258																																																																									
Db	908	GATCTTATGGTGAACACTTTAGGAATGCTGGCCAAATCAGGGCATATGTCGTATAC	967																																																																									
Qy	1259	GGCACTCATTTTTTGGACCATGAGATCAATCACTCTGTGGACCAAGATTTAAACCATTAAC	1318																																																																									
Db	968	GGCACTCATTTTTTGGACCATGAGATCAATCACTCTGTGGACCAAGATTTAAACCATTAAC	1027																																																																									
Qy	1319	CAGACTATGGCTCTCTCCCTCTCAAGCGGCTCATGGGCCCAAGCTTTGGCTGTC	1378																																																																									
Db	1028	CAGACTATGGCTCTCTCTCTCTCAAGCGGCTCATGGGCCCAAGCTTTGGCTGTC	1087																																																																									
Qy	1379	ATGTGGCTGGGCTCCAGCGGAAGCCAGGCTGTGCGAGTGATCCGGACAACCTAAGGA	1438																																																																									
Db	1088	ATGTGGCTGGGCTCCAGCGGAAGCCAGGCTGTGCGAGTGATCCGGACAACCTAAGGA	1147																																																																									
Qy	1439	TTTATGCTCATGTCACAACCAACCAACAACAACAACAACAACAACAACAACAACAACA	1498																																																																									
Db	1148	TTTATGCTCATGTCACAACCAACCAACAACAACAACAACAACAACAACAACAACAACA	1207																																																																									
Qy	1499	TTATCATCAACTTGCATGCATCAGAGAAAGAAATCAAGCTGTGGAGCTTCCAGAGACA	1558																																																																									
Db	1208	TTATCATCAACTTGCATGCATCAGAGAAAGAAATCAAGCTGTGGAGCTTCCAGAGACA	1267																																																																									
Qy	1559	AGCTGGTTCACCAAGTACCTGTGTCGACGCTTATGGGACAGAGGCTTAAGTCCAGTAC	1618																																																																									
Db	1258	AGCTGGTTCACCAAGTACCTGTGTCGACGCTTATGGGACAGAGGCTTAAGTCCAGTAC	1327																																																																									
Qy	1619	TGCAACTGATATGGCCAGCCCTTAGTGATGTGTGAGTCAGCGGACCTCCCAAGATTGAAGC	1678																																																																									
Db	1328	TGCAACTGATATGGCCAGCCCTTAGTGATGTGTGAGTCAGCGGACCTCCCAAGATTGAAGC	1387																																																																									
Qy	1679	CCCGGCCCTTGGGCGCGCGCGGACATTGTCATCCCTCCAGTCCAGTCCAGTGGGCTTTATG	1738																																																																									
Db	1388	CCCGGCCCTTGGGCGCGCGCGGACATTGTCATCCCTCCAGTCCAGTCCAGTGGGCTTTATG	1447																																																																									
Qy	1739	TGCTCAAGATGTCAATGCTTTGGCTGCTGCGCTGATCCAGTCAACATGGGCTTTATG	1779																																																																									
Db	1448	TGCTCAAGATGTCAATGCTTTGGCTGCTGCGCTGATCCAGTCAACATGGGCTTTATG	1488																																																																									
<div> <div>RESULT 7</div> <div>AAA91098</div> <div> <div>AAA91098 standard; DNA; 1898 BP.</div> <div> <div>AAA91098;</div> <div>AAA91098;</div> </div> </div> </div> <tr> <td>AC</td><td>AAA91098;</td><td></td><td></td></tr> <tr> <td>DT</td><td>20-APR-2001 (first entry)</td><td></td><td></td></tr> <tr> <td>DE</td><td>Human heparanase, hnhp1 p99 form, coding sequence.</td><td></td><td></td></tr> <tr> <td>XX</td><td>Heparanase; hnhp1; wound healing; angiogenesis; restenosis; Scurvy;</td><td></td><td></td></tr> <tr> <td>XX</td><td>atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;</td><td></td><td></td></tr> <tr> <td>XX</td><td>neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;</td><td></td><td></td></tr> <tr> <td>XX</td><td>gene therapy; human; ds.</td><td></td><td></td></tr> <tr> <td>OS</td><td>Homo sapiens.</td><td></td><td></td></tr> <tr> <td>XX</td><td></td><td></td><td></td></tr> <tr> <td>XX</td><td>Key</td><td>Location/Qualifiers</td><td></td></tr> <tr> <td>XX</td><td>CDS</td><td>26..1642</td><td></td></tr> <tr> <td>XX</td><td>FT</td><td>/tag= a</td><td></td></tr> <tr> <td>XX</td><td>FT</td><td>/product= "heparanase"</td><td></td></tr> <tr> <td>XX</td><td>FT</td><td>/transl_except= (pos:938..940, aa:Ala)</td><td></td></tr> <tr> <td>XX</td><td>PN</td><td>WO20010643-A2.</td><td></td></tr> <tr> <td>XX</td><td>PD</td><td>04-JAN-2001.</td><td></td></tr> <tr> <td>XX</td><td>PF</td><td>19-JUN-2000; 2000MO-IL000358.</td><td></td></tr> <tr> <td>XX</td><td>PR</td><td>25-JUN-1999; 99US-0140801P.</td><td></td></tr> <tr> <td>XX</td><td>PA</td><td>(INSI-) INSIGHT STRATEGY &amp; MARKETING LTD.</td><td></td></tr>	AC	AAA91098;			DT	20-APR-2001 (first entry)			DE	Human heparanase, hnhp1 p99 form, coding sequence.			XX	Heparanase; hnhp1; wound healing; angiogenesis; restenosis; Scurvy;			XX	atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;			XX	neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;			XX	gene therapy; human; ds.			OS	Homo sapiens.			XX				XX	Key	Location/Qualifiers		XX	CDS	26..1642		XX	FT	/tag= a		XX	FT	/product= "heparanase"		XX	FT	/transl_except= (pos:938..940, aa:Ala)		XX	PN	WO20010643-A2.		XX	PD	04-JAN-2001.		XX	PF	19-JUN-2000; 2000MO-IL000358.		XX	PR	25-JUN-1999; 99US-0140801P.		XX	PA	(INSI-) INSIGHT STRATEGY & MARKETING LTD.	
AC	AAA91098;																																																																											
DT	20-APR-2001 (first entry)																																																																											
DE	Human heparanase, hnhp1 p99 form, coding sequence.																																																																											
XX	Heparanase; hnhp1; wound healing; angiogenesis; restenosis; Scurvy;																																																																											
XX	atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;																																																																											
XX	neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;																																																																											
XX	gene therapy; human; ds.																																																																											
OS	Homo sapiens.																																																																											
XX																																																																												
XX	Key	Location/Qualifiers																																																																										
XX	CDS	26..1642																																																																										
XX	FT	/tag= a																																																																										
XX	FT	/product= "heparanase"																																																																										
XX	FT	/transl_except= (pos:938..940, aa:Ala)																																																																										
XX	PN	WO20010643-A2.																																																																										
XX	PD	04-JAN-2001.																																																																										
XX	PF	19-JUN-2000; 2000MO-IL000358.																																																																										
XX	PR	25-JUN-1999; 99US-0140801P.																																																																										
XX	PA	(INSI-) INSIGHT STRATEGY & MARKETING LTD.																																																																										

XX	Packer I, Michal I, Itzhaki H;
PI	
XX	WPI, 2001-137930/14.
DR	P-PSDB; AAY97633.
XX	
PT	New polynucleotides and polypeptides that are distantly homologous to
PT	heparanase, useful in wound healing, as well as in gene therapy protocols
PT	for angiogenesis, restenosis, atherosclerosis, or inflammation.
XX	
PS	Claim 3; Page 60-61; 67pp; English.
XX	
CC	This sequence encodes a heparanase of the invention. The heparanase DNA
CC	and protein sequences are useful in wound healing, angiogenesis,
CC	restenosis, atherosclerosis, inflammation, pulmonary diseases, and
CC	neurodegenerative diseases (such as Scrapie, Alzheimer's disease, and
CC	Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
CC	sequence is particularly useful in gene therapy
XX	
SQ	Sequence 1898 BP; 470 A; 523 C; 457 G; 448 T; 0 U; 0 Other;
	Query Match 81.1%; Score 1443.4; DB 4; Length 1898;
	Best Local Similarity 90.8%; Pred. No. 0;
	Matches 1616; Conservative 0; Mismatches 1; Indels 162; Gaps 1;
OY	1 ATGAGGGTGCCTTTGTGCTTCCTCCCGAAGCATTGCCCTCCAGCAATCCCGCCCCCGCG 60
D6	26 ATGAGGGTGCCTTTGTGCTTCCTCCCGAAGCATTGCCCTCCAGCAATCCCGCCCCCGCG 85
OY	61 TGCTTAGCCCCGGGGGCTCTCTACTTGCTGTGCTCTCATCTCTCTCTCTCCACG 120
D6	86 TGCTTAGCCCCGGGGGCTCTCTACTTGCTGTGCTCTCATCTCTCTCTCTCTCCACG 145
OY	121 GCTGGAACAAGAACCCTTGCTGCTGAGACAGAGCTCAGGTTGAAAGAAAACCTGTG 180
D6	146 GCTGGAACAAGAACCCTTGCTGCTGAGACAGAGCTCAGGTTGAAAGAAAACCTGTG 205
OY	181 ATTCTACTGATGTAGAACCAAGAACCCAGTCAGACAGTCAATGAGAACTTCTCTCT 240
D6	206 ATTCTACTGATGTAGAACCAAGAACCCAGTCAGACAGTCAATGAGAACTTCTCTCT 265
OY	241 CTGCAAGCTGATCCGTCATCATTCATGATGAGGCTGCTCATTTCTCTAAGCTCCAAGCGC 300
D6	266 CTGCAAGCTGATCCGTCATCATTCATGATGAGGCTGCTCATTTCTCTAAGCTCCAAGCGC 325
OY	301 TTGTGACCTCGGCCGGGGACTTTGCGCCGCTTTCTGCGCTTGGGGGCAAAAAGAAC 360
D6	326 TTGTGACCTCGGCCGGGGACTTTGCGCCGCTTTCTGCGCTTGGGGGCAAAAAGAAC 385
OY	361 GACTTCTGCAAGTTCAGAAACCTGAGGAAACCGGCGAAAAAGCCGCGGGGCGCCG 420
D6	386 GACTTCTGCAAGTTCAGAAACCTGAGGAAACCGGCGAAAAAGCCGCGGGGCGCCG 445
OY	421 GATTACTATCTCAAAAACCTATGAGATGACATTGTTGAAAGTAATGTTGCTTAGATAAA 480
D6	446 GATTACTATCTCAAAAACCTATGAGAT----- 472
OY	481 CAGAAAGGCTGCAAGATTGGCCAGACCTCGATGTATGCTGTGCTCCAAAGGAGAG 540
D6	473 ----- 472
OY	541 GCAGCTCAGATGCATCTGTGTTCTTTAAAGGAGCAATTCCTCAAATCTTAAGTAATCTC 600
D6	473 ----- 472
OY	601 ATATTAAAGCAGAGTCTAGACAAACTTATTAATTTGCTGATGCTCTGAGCTCCAC 660
D6	473 -----GCAGAGTCTTAGACAAACTTATTAATTTGCTGATGCTCTGAGCTCCAC 523
OY	661 CTGATATTTGCTTAATGCACTGCGTCGTAAATCCCAATACTCTCTGGAAGAAGTCTTAGT 720
D6	524 CTGATATTTGCTTAATGCACTGCGTCGTAAATCCCAATACTCTCTGGAAGAAGTCTTAGT 583

[illegible]

Query Match	Best Local Similarity	Matches 1614; Conservative	Score 1441.4; DB 6; Length 1685; Fred. No. 0; Mismatches 2; Indels 162; Gaps 1
1	ATGAGGCTGCTTTGTCCTTCCTGAGGCAATGCTTCGACCAACTCCGCCCCCGCG	60	
55	ATGAGGCTGCTTTGTCCTTCCTGAGGCAATGCTTCGACCAACTCCGCCCCCGCG	114	
61	TGCTGACGCCCCGGGGGCTCTCTACTTGGCTCTGTGTGTCATGATCTTCCTTCCTCCAG	120	
115	TGCTGACGCCCCGGGGGCTCTCTACTTGGCTCTGTGTGTCATGATCTTCCTTCCTCCAG	174	

QY 121 GCTGAGACAGGAGACCTTGGCTCTGTAGACAGAGCTGCAGGTTTGAAGAAAAGACCTG 180  
 DB 175 GCTGAGAGACAGAGACCTTGGCTCTGTAGACAGAGCTGCAGGTTTGAAGAAAAGACCTG 234  
 QY 181 ATTCTACTGATGTAGTGTAGACACCAAGAACCCAGTCAAGACAGTCAATAGAGACTTCTCT 240  
 DB 235 ATTCTACTGATGTAGTGTAGACACCAAGAACCCAGTCAAGACAGTCAATAGAGACTTCTCT 294  
 QY 241 CTGACAGTGTATCTGCTCATCTATCATGTATGTATGTATGTATGTATGTATGTATGTAT 300  
 DB 295 CTGACAGTGTATCTGCTCATCTATCATGTATGTATGTATGTATGTATGTATGTATGTAT 354  
 QY 301 TTGGTGAACCTGGGCGCGGAGACCTTGGCGGCTTCTGCGGCTTGGGCGGCGGCGGCGG 360  
 DB 355 TTGGTGAACCTGGGCGCGGAGACCTTGGCGGCTTCTGCGGCTTGGGCGGCGGCGGCGG 414  
 QY 361 GACTTCCTGAGTTCAGAACCTGAGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
 DB 415 GACTTCCTGAGTTCAGAACCTGAGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474  
 QY 421 GATTACTATCTCAAAAACCTATGTAGATGACATTTGTTCAGAGATGATGATGATGATGAT 480  
 DB 475 GATTACTATCTCAAAAACCTATGTAGATGACATTTGTTCAGAGATGATGATGATGATGAT 501  
 QY 481 CAGAAAGGCTGCAAGATTTGCCAGACCCCTGATGTTATGCTGCTCCAAAGGAGAG 540  
 DB 502 ----- 501  
 QY 541 GCAGCTCAGATGATCTGTTCTTTTAAAGAGCAATTTCCAAATCTTAACATATCTC 600  
 DB 502 ----- 501  
 QY 601 ATATTAACAGCCAGTCTCTAGACAACTTTATTAATCTTGTGATGCTGTGACTCCAC 660  
 DB 502 -----GCCAGGCTCTAGACAACTTTATTAATCTTGTGATGCTGTGACTCCAC 552  
 QY 661 CTGATATTTGCTCTTAATGTAGCTGCTGTATGCTCAATTAATCTCTGGAACGTTTACT 720  
 DB 553 CTGATATTTGCTCTTAATGTAGCTGCTGTATGCTCAATTAATCTCTGGAACGTTTACT 612  
 QY 721 GCCCTAGTCTGTTGAAGTACAGGCGCAGCAAAATTTCAACAATTTCTGGGAGCTGGT 780  
 DB 613 GCCCTAGTCTGTTGAAGTACAGGCGCAGCAAAATTTCAACAATTTCTGGGAGCTGGT 672  
 QY 781 AATGAGCAAAATTAATCTATCGACCATGATGCGCGGCGAGTAAATGSCAGCAGTTGGA 840  
 DB 673 AATGAGCAAAATTAATCTATCGACCATGATGCGCGGCGAGTAAATGSCAGCAGTTGGA 732  
 QY 841 AAGGATTAATCTATCGACCATGATGCGCGGCGAGTAAATGSCAGCAGTTGGA 900  
 DB 733 AAGGATTAATCTATCGACCATGATGCGCGGCGAGTAAATGSCAGCAGTTGGA 792  
 QY 901 TTATATGCGCTTAATTTGGGCGCGGAGGAAATGATGATGCTCTCTAGATGATTC 960  
 DB 793 TTATATGCGCTTAATTTGGGCGCGGAGGAAATGATGATGCTCTCTAGATGATTC 852  
 QY 961 ATGAGGTGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 853 ATGAGGTGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912  
 QY 1021 CGGAGTGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 913 CGGAGTGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972  
 QY 1081 ATTGAGAAATTCAGAAAGTGTATTAATCAATCACTCCAGGAAAGAAATTTGGCTTGA 1140  
 DB 973 ATTGAGAAATTCAGAAAGTGTATTAATCAATCACTCCAGGAAAGAAATTTGGCTTGA 1032  
 QY 1141 GGATGTGTGACACCTCAGCTGAGGAGCAAAATATATCCGATTCCTATGCTGAGGA 1200  
 DB 1033 GGATGTGTGACACCTCAGCTGAGGAGCAAAATATATCCGATTCCTATGCTGAGGA 1092

QY 1201 TTCTATGTTGAACACTTAGAATGCTGCGCAATCAGGCGCATGATGATGATGATGATGATGAT 1260  
 DB 1093 TTCTATGTTGAACACTTAGAATGCTGCGCAATCAGGCGCATGATGATGATGATGATGATGAT 1152  
 QY 1261 CACTGATTTTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 1153 CACTGATTTTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212  
 QY 1321 GACTACTGCT 1380  
 DB 1213 GACTACTGCT 1272  
 QY 1381 GTGCTGAGGCTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1440  
 DB 1273 GTGCTGAGGCTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1332  
 QY 1441 TATGCTGATGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 DB 1333 TATGCTGATGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1392  
 QY 1501 ATCATCACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 DB 1393 ATCATCACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1452  
 QY 1561 CTGCTTCAACAGTACCTGCTGACGCTTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
 DB 1453 CTGCTTCAACAGTACCTGCTGACGCTTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1512  
 QY 1621 CACTGATGCGGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB 1513 CACTGATGCGGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572  
 QY 1681 CGCCCGCTTGGGCGGCGGCGGAGTGTATGCTATGCTTCAATGATGATGATGATGATGAT 1740  
 DB 1573 CGCCCGCTTGGGCGGCGGCGGAGTGTATGCTATGCTTCAATGATGATGATGATGATGAT 1632  
 QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 DB 1633 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1671

RESULT 9  
 AB211527  
 ID AB211527 standard; cDNA; 1891 BP.  
 XX  
 AC AB211527;  
 XX  
 DT 20-JAN-2003 (first entry)  
 XX  
 DE Human polynucleotide seq ID NO 409.  
 XX  
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antichrinitic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200270539-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 05-MAR-2002; 2002WO-US005095.  
 XX  
 PR 05-MAR-2001; 2001US-00799451.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;  
 DR WPI; 2002-759812/82.  
 XX P-PSDB; ABP69310.

PT New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.

XX Claim 1; SEQ ID NO 409; 1012bp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence selected from any of 948 sequences (AB21119-  
 CC AB21066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
 CC arthritis, etc. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 1891 BP; 483 A; 513 C; 456 G; 439 T; 0 U; 0 Other;

Query Match 79.7%; Score 1417.8; DB 6; Length 1891;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

QY 1 ATGAGGGTGGTGTGCTTCCCTGAGGAGCATGCTCCAGCAACTCCCGCCCGCG 60  
 DB 20 ATGAGGGTGGTGTGCTTCCCTGAGGAGCATGCTCCAGCAACTCCCGCCCGCG 79  
 QY 61 TGGCTTACCCCGGCGGCTCTTACTTGGCTCTGCTCCATCTCTCCCTTCTCCAG 120  
 DB 80 TGGCTTACCCCGGCGGCTCTTACTTGGCTCTGCTCCATCTCTCCCTTCTCCAG 139  
 QY 121 GCGGAGACGAGAGACCTTGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 DB 140 GCGGAGACGAGAGACCTTGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199  
 QY 181 ATTCTACTGATGTGAGACCAAGAACCCAGTCAGAGCATGATGAGAACTTCTCT 240  
 DB 200 ATTCTACTGATGTGAGACCAAGAACCCAGTCAGAGCATGATGAGAACTTCTCT 259  
 QY 241 CTGAGCTGATCCCTGATCATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 260 CTGAGCTGATCCCTGATCATGATGATGATGATGATGATGATGATGATGATGAT 319  
 QY 301 TTGGTGAACCTTGGCCCGGAGACTTGGCCCGCTTCTGCGCTTCCGCGGCAAGAGAC 360  
 DB 320 TTGGTGAACCTTGGCCCGGAGACTTGGCCCGCTTCTGCGCTTCCGCGGCAAGAGAC 379  
 QY 361 GACTTCTGAGATTCGAGAACTTGAAGAACCCGCGGCGGCGGCGGCGGCGGCGG 420  
 DB 380 GACTTCTGAGATTCGAGAACTTGAAGAACCCGCGGCGGCGGCGGCGGCGGCGG 439  
 QY 421 GATTACTATCTCAAAAATCTAGAGATGATGATGATGATGATGATGATGATGATGAT 480  
 DB 440 GATTACTATCTCAAAAATCTAGAGATGATGATGATGATGATGATGATGATGATGAT 499  
 QY 481 CAGAAAGGCTGCAAGATTCGCGGAGACCTGATGATGATGATGATGATGATGATGATGAT 540  
 DB 500 CAGAAAGGCTGCAAGATTCGCGGAGACCTGATGATGATGATGATGATGATGATGATGAT 559  
 QY 541 GCAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

DB 560 GCAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 619  
 QY 601 ATATTAACAGCCAGGCTCTGAGCAAACTTTAATCTTGTGATGCTGAGACCCAGC 660  
 DB 620 ATATTAACA----- 628  
 QY 661 CTGATATTTGCTCTAAATGCACTGCGTGAATCCCAATPACTCTGAGACAGTTCTAGT 720  
 DB 629 ----- 628  
 QY 721 GCCCTGAGTCTGTTGAAGTACAGCCGCAAGCAAAAGTACAACTTTCTTGGGAACCTGGGT 780  
 DB 629 ----- 628  
 QY 781 AATGAGCCAAATATCTATCGGACCATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB 629 ---GAGCCAAATATCTATCGGACCATGATGATGATGATGATGATGATGATGATGATGAT 685  
 QY 841 AAGGATTAATCTCAAGCTGAAAGAGCTGTTGACCCCATCCGATTTATTCAGAGCCAGC 900  
 DB 686 AAGGATTAATCTCAAGCTGAAAGAGCTGTTGACCCCATCCGATTTATTCAGAGCCAGC 745  
 QY 901 TTATATGACCCCTATATTTGGAGCGGCGGAGAGATGATGATGATGATGATGATGATGAT 960  
 DB 746 TTATATGACCCCTATATTTGGAGCGGCGGAGAGATGATGATGATGATGATGATGATGAT 805  
 QY 961 ATGAGAGTGGCAGGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 806 ATGAGAGTGGCAGGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGAT 865  
 QY 1021 CCGGTGTGTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 866 CCGGTGTGTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925  
 QY 1081 ATTAGAAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGAAATTTGCTTGA 1140  
 DB 926 ATTAGAAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGAAATTTGCTTGA 985  
 QY 1141 GGTGTGTGACCACTCACTGATGAGAGCAAAACATATCCGATTCCTGATGAGAGAG 1200  
 DB 986 GGTGTGTGACCACTCACTGATGAGAGCAAAACATATCCGATTCCTGATGAGAGAG 1045  
 QY 1201 TTTTATGATGTAACAATTTAGAGATGCTGCGCAATCAGGAGATGATGATGATGATGAT 1260  
 DB 1046 TTTTATGATGTAACAATTTAGAGATGCTGCGCAATCAGGAGATGATGATGATGATGAT 1105  
 QY 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 1106 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165  
 QY 1321 GACTACTGCTCTCTCTCTCTCAACAAGGCTGATGCGGCCCAAGTCTTGTGCTGAT 1380  
 DB 1166 GACTACTGCTCTCTCTCTCTCAACAAGGCTGATGCGGCCCAAGTCTTGTGCTGAT 1225  
 QY 1381 GTGGCTGGGCTCCAGCGGAAGCCAGGCTGCGGAGATTCGCGGCAAACTTAAGAT 1440  
 DB 1226 GTGGCTGGGCTCCAGCGGAAGCCAGGCTGCGGAGATTCGCGGCAAACTTAAGAT 1285  
 QY 1441 TATGCTCACTGCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACA 1500  
 DB 1286 TATGCTCACTGCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACA 1345  
 QY 1501 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 DB 1346 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1405  
 QY 1561 CTGCTTCAAGTACCTGCTGACGCTTATGAGGAGAGGCTTAAAGTCCAAATGCAAGT 1620  
 DB 1406 CTGCTTCAAGTACCTGCTGACGCTTATGAGGAGAGGCTTAAAGTCCAAATGCAAGT 1465  
 QY 1621 CAAGTAAAGGCGACCCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

Db 1466 CAACGTGAATGGCCAGCCCTTACTGATGTGAGACGACGAGCCCTCCAGATTGAAGCCC 1525  
 Qy 1661 GCGCCCTTTGGGCGCGCGGCGGACATTTGTCATCCCTCCAGTCACCATGGGCTTTTATGTG 1740  
 Db 1526 GCGCCCTTTGGGCGCGCGGCGGACATTTGTCATCCCTCCAGTCACCATGGGCTTTTATGTG 1585  
 Qy 1741 GTCAAGATGTCATGCTTTTGGCTGCGCTACCGATTA 1779  
 Db 1586 GTCAAGATGTCATGCTTTTGGCTGCGCTACCGATTA 1624

RESULT 10  
 AA170705  
 ID AA170705 standard; cDNA; 2326 BP.  
 AC AA170705;  
 XX 04-FEB-2002 (first entry)  
 XX  
 XX Human prepro-heparanase II cDNA.  
 DE Heparanase II; human; vulnery; angiogenesis inhibitor;  
 KM antinflammatory; cyrostatic; gene therapy; diagnosis; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 25..1629  
 XX FT /tag= a  
 XX FT 25..147  
 XX FT sig\_peptide /tag= b  
 XX FT 148..1626  
 XX FT mat\_peptide /tag= c  
 XX  
 XX MO200181569-A2.  
 XX  
 XX 01-NOV-2001.  
 XX  
 XX 17-APR-2001; 2001MO-US010804.  
 XX  
 XX 20-APR-2000; 2000US-0199072P.  
 XX  
 XX (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 XX Heinrikson RL, Bienkowski MJ;  
 XX WPI; 2002-041402/05.  
 XX P-PSDB; AAM50337.  
 XX  
 XX Novel heparanase II polypeptide useful for identifying agents with alter  
 PT heparanase activity and for accelerating wound healing, blocking  
 PT angiogenesis or inflammation.  
 PS  
 PS Claim 2; Page 9-11; 65pp; English.  
 XX  
 CC The present sequence is that of cDNA encoding novel human prepro-  
 CC heparanase II (see AAM50337), a paralogue of human heparanase I.  
 CC Identification of the cDNA involved a combination of protein sequencing  
 CC and mining of expressed sequence tag databases using human heparanase I  
 CC as query. Incyte clones 1634352 (prostate tumour), 3207353 (corpus  
 CC cavernosum) and 3704980 (corpus cavernosum) were detected, and additional  
 CC sequences were obtained from Incyte clones 3529440 (bladder) and 3385825  
 CC (oesophagus) and from 5' RACE products using human prostate, small  
 CC intestine, bladder or heart cDNA templates. Heparanase II is a secreted  
 CC protein that shows 43% identity at the amino acid level to heparanase I.  
 CC The prepro-protein encoded by the current sequence is processed to remove  
 CC a 41-amino acid leader peptide, and further processed to remove internal  
 CC amino acids, yielding the 8 kDa and 50 kDa subunits of the heparanase II  
 CC enzyme. Heparanase I and II have a non-overlapping expression pattern in  
 CC human tissues and each may serve tissue-specific functional roles. The  
 CC invention provides heparanase II nucleic acids, vectors, host cells,  
 CC polypeptides and antibodies. Heparanase II is useful for identifying an  
 CC agent that alters heparanase activity, and which can be used to treat a

CC disease state. Nucleic acids are useful for large scale expression of  
 CC heparanase II, in the design of probes and primers, and in the  
 CC identification of a genetic alteration(s) that underlies a disease state or  
 CC states. Inhibiting heparanase II activity is useful for treating or  
 CC preventing metastasis, cancer, CNS and neurodegenerative diseases,  
 CC inflammation and cardiovascular disease such as restenosis following  
 CC angioplasty and atherosclerosis. Heparanase II is useful for accelerating  
 CC wound healing, blocking angiogenesis, degradation of heparin and  
 CC neutralization of heparin's anticoagulant properties during surgery.  
 CC Heparanase or an agent that enhances heparanase activity can also be  
 CC infused into the vasculature to block accumulation and diapedesis of  
 CC neutrophils at sites of inflammation with or without added domains to  
 CC confer selectivity in delivery

Sequence 2326 BP; 593 A; 631 C; 550 G; 552 T; 0 U; 0 Other;  
 XX  
 XX  
 XX Query Match 79.7%; Score 1417.8; DB 6; Length 2326;  
 XX Best Local Similarity 90.1%; Pred. No. 0;  
 XX Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

Qy 1 ATGAGGCTGCTTTTGTGCTTCTCTGAGCATGCTCCAGCAACCCGCCCCCGCG 60  
 Db 25 ATGAGGCTGCTTTTGTGCTTCTCTGAGCATGCTCCAGCAACCCGCCCCCGCG 84  
 Qy 61 TGCCCTAGCCCCGGGGGCTCTCTACTTGAGCTCTGCTGCTCATCTCCCTTCTCCAG 120  
 Db 85 TGCCCTAGCCCCGGGGGCTCTCTACTTGAGCTCTGCTGCTCATCTCCCTTCTCCAG 144  
 Qy 121 GCTGAGACAGAGAGACCTTGTGCTGTAGACAGAGCTTGACAGTGAAGAAAGACCTG 180  
 Db 145 GCTGAGACAGAGAGACCTTGTGCTGTAGACAGAGCTTGACAGTGAAGAAAGACCTG 204  
 Qy 181 ATTCTACTGTATGTGACACCAAGAACCCAGTCAGAGAGCTCAATGAGAACTTCTCT 240  
 Db 205 ATTCTACTGTATGTGACACCAAGAACCCAGTCAGAGAGCTCAATGAGAACTTCTCT 264  
 Qy 241 CTGCAAGCTGATCCGTCATCATTCATGATGCTGCTGCTGATTTCTTAAGCTCCAAAGCG 300  
 Db 265 CTGCAAGCTGATCCGTCATCATTCATGATGCTGCTGCTGATTTCTTAAGCTCCAAAGCG 324  
 Qy 301 TTGGTGAACCTGAGCGCGGAGACTTGGCCGCTTTTGTGGCTTCGGGGGCAAAAGAGC 360  
 Db 325 TTGGTGAACCTGAGCGCGGAGACTTGGCCGCTTTTGTGGCTTCGGGGGCAAAAGAGC 384  
 Qy 361 GACTTCCGTCAGTTCAGAAACCTGAGAGACCCGCGCAAAAGCCGCGGGGCGCG 420  
 Db 385 GACTTCCGTCAGTTCAGAAACCTGAGAGACCCGCGCAAAAGCCGCGGGGCGCG 444  
 Qy 421 GATTACTATCTCAAAAACCTATGAGAGATGATTTGGAAGTATTTGCTTATGATTA 480  
 Db 445 GATTACTATCTCAAAAACCTATGAGAGATGATTTGGAAGTATTTGCTTATGATTA 504  
 Qy 481 CAGAAAGCTGCAAGATGAGCCAGACCGCTGATGTATGCTGAGCTCCAAAGGAGAG 540  
 Db 505 CAGAAAGCTGCAAGATGAGCCAGACCGCTGATGTATGCTGAGCTCCAAAGGAGAG 564  
 Qy 541 GCGCTCGATGATGATCTGCTTTCTTAAAGAGCAATTTCTCAATCTTACAGTAATCTC 600  
 Db 565 GCGCTCGATGATGATCTGCTTTCTTAAAGAGCAATTTCTCAATCTTACAGTAATCTC 624  
 Qy 601 ATATTAAGCAAGCTCTGAGCAAACTTATTAATCTTGTATGCTCTGAGACTCAC 660  
 Db 625 ATATTAAGCAAGCTCTGAGCAAACTTATTAATCTTGTATGCTCTGAGACTCAC 684  
 Qy 661 CTGATATTTGCTTAATGACATGCGGTGATCCCAATTACTCTGAGCAATTTAGT 720  
 Db 684 CTGATATTTGCTTAATGACATGCGGTGATCCCAATTACTCTGAGCAATTTAGT 744  
 Qy 721 GCGCTGAGTCTGTGAATGACAGCGCCAGCAAAAGTACACATTTCTTGGAACTGGGT 780  
 Db 744 GCGCTGAGTCTGTGAATGACAGCGCCAGCAAAAGTACACATTTCTTGGAACTGGGT 804  
 Qy 781 AATGAGCAAAATTAATCTATGAGCAATGATGCGCGGAGATTAATGAGCAAGTGGGA 840  
 Db 804 AATGAGCAAAATTAATCTATGAGCAATGATGCGCGGAGATTAATGAGCAAGTGGGA 868



```

Db 634 ---GAGCAATATACATACGACATGCGACGCGGACAGTAATAGCAGCCAGTTGGGA 690
QY 841 AAGATTACATCCAGCGAGAGCCTGTTGACGCCATCCGAGTTTATTCAGAGCCAGC 900
Db 691 AAGGATTACATCCAGCGAGAGCCTGTTGACGCCATCCGAGTTTATTCAGAGCCAGC 750
QY 901 TTATATGCCCCATATATATGCGCGCCGAGAGAGATGTCATCGCCCTCTAGATGATTC 960
Db 751 TTATATGCCCCATATATATGCGCGCCGAGAGAGATGTCATCGCCCTCTAGATGATTC 810
QY 961 ATGAGGTGCGAGAGATACAGTATGATGCGATGATACCTGCGAATATGCTTATGATGCG 1020
Db 811 ATGAGGTGCGAGAGATACAGTATGATGCGATGATACCTGCGAATATGCTTATGATGCG 870
QY 1021 CGGGTGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 871 CGGGTGTCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 930
QY 1081 ATTAGGAAATTCAGAAAGTGTGTTATATACATACATTCAGAGAAAGAAATTTGGCTTGA 1140
Db 931 ATTAGGAAATTCAGAAAGTGTGTTATATACATACATTCAGAGAAAGAAATTTGGCTTGA 990
QY 1141 GGTGTGTGACACACCTGAGTGAAGGACAAACATCTATCCATTCCTATGCTGACAGA 1200
Db 991 GGTGTGTGACACACCTGAGTGAAGGACAAACATCTATCCATTCCTATGCTGACAGA 1050
QY 1201 TTCTTATGTTGAAACACTTTAGAAATGCTGCGCAATGAGGACATGATGCTGATACG 1260
Db 1051 TTCTTATGTTGAAACACTTTAGAAATGCTGCGCAATGAGGACATGATGCTGATACG 1110
QY 1261 CACTCATTTTTTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1111 CACTCATTTTTTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1170
QY 1321 GACTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Db 1171 GACTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1230
QY 1381 GTGCTGTGGGCTCCGAGCGGAGCCAGCGGCTGCGAGTATCCGGAGCAAACTTAAGATT 1440
Db 1231 GTGCTGTGGGCTCCGAGCGGAGCCAGCGGCTGCGAGTATCCGGAGCAAACTTAAGATT 1290
QY 1441 TATGCTCACTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCACTTT 1500
Db 1291 TATGCTCACTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCACTTT 1350
QY 1501 ATCATCAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db 1351 ATCATCAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1410
QY 1561 CTGCTTACACAGTACCTGCTGACGCCCTATGCGAGAGAGGCGCTTAAGTCCAGTACAGTG 1620
Db 1411 CTGCTTACACAGTACCTGCTGACGCCCTATGCGAGAGAGGCGCTTAAGTCCAGTACAGTG 1470
QY 1621 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db 1471 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1530
QY 1681 CGCCCCCTTCCGAGCGCGAGCACTTGTGATGATGATGATGATGATGATGATGATGATG 1740
Db 1531 CGCCCCCTTCCGAGCGCGAGCACTTGTGATGATGATGATGATGATGATGATGATGATG 1590
QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
Db 1591 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1629

```

## RESULT 11

AAH22672  
ID AAH22672 standard; DNA, 2462 BP.  
XX AC AAH22672;

```

XX 07-SEP-2001 (first entry)
DT Heparanase-like protein Hpa2 splice variant #2 encoding DNA.
XX DE
XX KW Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;
XX KW cytosolic; neuroprotective; cerebroprotective; immunosuppressive;
XX KW antiproliferative; neurotrophic; antiinflammatory; antiarthritic; antidiabetic;
XX KW antidiabetic; antiarteriosclerotic; vulnary; de.
OS Homo sapiens.
XX FH
XX FH Key Location/Qualifiers
XX FT 5'UTR 1..600
XX FT CDS /*tag= a
XX FT 601..2205 /*tag= b
XX FT 3'UTR 2206..2462 /*tag= c
XX FT
XX WO200146392-A2.
XX PN
XX PD 28-JUN-2001.
XX PF 21-DEC-2000; 2000WO-GB004963.
XX PR 22-DEC-1999; 99GB-00030392.
XX PR 07-APR-2000; 2000GB-00008713.
XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX PI McKenzie EA, Stamps AC, Terrett JA, Tyson KL;
XX DR MPI. 2001-418056/44.
XX DR P-PSDB; AAB85216.
XX PT Novel homologs of heparanase, present in three splice variants, useful
XX PT for identifying agents that modulate heparanase, useful in the treatment
XX PT and/or prophylaxis of abnormal levels of heparanase.
XX PS Claim 12; Fig 2; 97pp; English.
XX CC The invention provides a homologue to heparanase which is present in
XX CC three splice variants. The heparanase homologue polypeptide is useful in
XX CC the treatment of a human or non-human animal or for use in diagnosis.
XX CC Vectors comprising the heparanase homologue polynucleotides are useful in
XX CC the transformation or transfection of a prokaryotic or eukaryotic host.
XX CC The modulators of the polypeptide are useful in the manufacture of a
XX CC medicament for the treatment and/or prophylaxis of a condition/disease
XX CC associated with abnormal levels of the heparanase homologue, including
XX CC cancer, central nervous system (CNS) and neurodegenerative diseases,
XX CC cardiovascular diseases such as restenosis following angioplasty and
XX CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,
XX CC arthritis, inflammatory diseases, arthritis, vascular restenosis, tumour
XX CC growth and progression, asthma, Alzheimer's disease, diabetic
XX CC retinopathy, wound healing and inflammation. The polypeptide is also
XX CC useful in diagnosis and research. The present sequence represents the
XX CC nucleotide sequence of the mid-sized splice variant of the heparanase-
XX CC like protein Hpa2 of the invention
SQ Sequence 2462 BP; 661 A; 621 C; 638 G; 541 T; 0 U; 1 Other;
Query Match 79.7%; Score 1417.8; DB 4; Length 2462;
Best Local Similarity 90.1%; Freq. No. 0;
Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;
QY 1 ATGAGGTGCTTTTGTGCTTCTCTGAGCGCATGCGCTTCAGCAACTCCGCGCGCGCG 60
Db 601 ATGAGGTGCTTTTGTGCTTCTCTGAGCGCATGCGCTTCAGCAACTCCGCGCGCGCG 660
QY 61 TGCCTAGCCCGGCGGCGCTCTACTTGGCTGTGTTGCTCCATCTCTCCCTTCTCCAG 120
Db 661 TGCCTAGCCCGGCGGCGCTCTACTTGGCTGTGTTGCTCCATCTCTCCCTTCTCCAG 720

```



```

QY 121 GCTGAGACAGAGAGACCTTGCTGCTGTAGACAGAGCTGACGAGTTTGAAGGAAAAGACCTTG 180
Db 721 GCTGAGACAGAGAGACCTTGCTGCTGTAGACAGAGCTGACGAGTTTGAAGGAAAAGACCTTG 780
QY 181 ATTCTACTGTGATGTGAGACCAAGAAACCAAGTCAGAGACAGTCAATGAGAACTTCTCTCT 240
Db 781 ATTCTACTGTGATGTGAGACCAAGAAACCAAGTCAGAGACAGTCAATGAGAACTTCTCTCT 840
QY 241 CTGACAGCTGGATCCGTCATCATTCATGATGAGCTGGCTGATTTCTGTAGCTCAAGGCG 300
Db 841 CTGACAGCTGGATCCGTCATCATTCATGATGAGCTGGCTGATTTCTGTAGCTCAAGGCG 900
QY 301 TTGGTACCTTGAGCCCGGAGACTTTTGCCGCTTCTGCGCTTGCGGGGCAAAAGAGAC 360
Db 901 TTGGTACCTTGAGCCCGGAGACTTTTGCCGCTTCTGCGCTTGCGGGGCAAAAGAGAC 960
QY 361 GACTTCTGCAAGTTCAGAACTGAGAAACCCGCGGAAAGCCGCGGGGCGCCGCGCG 420
Db 961 GACTTCTGCAAGTTCAGAACTGAGAAACCCGCGGAAAGCCGCGGGGCGCCGCGCG 1020
QY 421 GATTACTATCTCAAAACTATGAGATGACATTTGTCAGAGTGAATTTGCTTAGATAA 480
Db 1021 GATTACTATCTCAAAACTATGAGATGACATTTGTCAGAGTGAATTTGCTTAGATAA 1080
QY 481 CAGAAAGCTGCAAGATTGCCGACACCTGATGTTATGCTGTGCTCCAAAGGAGAG 540
Db 1081 CAGAAAGCTGCAAGATTGCCGACACCTGATGTTATGCTGTGAGCTCCAAAGGAGAG 1140
QY 541 GCAGCTCAGATGATCGGTTCTTTCTAAAGAGACATTTCCCAATCTACAGTAATCTC 600
Db 1141 GCAGCTCAGATGATCGGTTCTTTCTAAAGAGACATTTCCCAATCTACAGTAATCTC 1200
QY 601 AATATTAACGACGAGTCTCTAGACAAACTTATTAATCTTGTGATGCTGTGACTCAC 660
Db 1201 AATATTAACA----- 1209
QY 661 CTGATATTTGCTCTAATGACATGCGGTCCGTAATCCCAATACCTCTGAGACAGTTCTACT 720
Db 1210 ----- 1209
QY 721 GCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAGTACAACTTTCTTGGAACCTGAGT 780
Db 1210 ----- 1209
QY 781 AATGACCCAAATTAATCTATGAGACCATGATGCGCGGCAATTAATGCGACGAGTTGGA 840
Db 1210 ---GACCCAAATTAATCTATGAGACCATGATGCGCGGCAATTAATGCGACGAGTTGGA 1266
QY 841 AAGGATTACATCCAGCTGAAGAGCTGTTGACAGCCCATCCGAGTTTATTCAGAGCCAGC 900
Db 1267 AAGGATTACATCCAGCTGAAGAGCTGTTGACAGCCCATCCGAGTTTATTCAGAGCCAGC 1326
QY 901 TTATATGCGCTTAATATTTGGGCGCGCGAGAGAAATGTCATGCGCTCTCTAGATGATTC 960
Db 1327 TTATATGCGCTTAATATTTGGGCGCGCGAGAGAAATGTCATGCGCTCTCTAGATGATTC 1386
QY 961 ATGAGGTGCGCAGAGATGACAGTAGATGACAGTACCTGCGGCAATGCTGCTACATTAATG 1020
Db 1387 ATGAGGTGCGCAGAGATGACAGTAGATGACAGTACCTGCGGCAATGCTGCTACATTAATG 1446
QY 1021 CCGGTTGCTCAAGGTGATGACATTTCTGAAAACCTGCGCTGTAGACACACTCTCTGACAG 1080
Db 1447 CCGGTTGCTCAAGGTGATGACATTTCTGAAAACCTGCGCTGTAGACACACTCTCTGACAG 1506
QY 1081 ATTAGGAAAATTCAGAAAGTGTATTAATACATCACTCCAGGAAAGAAATTTGGCTTGA 1140
Db 1507 ATTAGGAAAATTCAGAAAGTGTATTAATACATCACTCCAGGAAAGAAATTTGGCTTGA 1566
QY 1141 GGTGTGTGTGACCACTCAGCTGAGGCAAAACATCTATCCGATTCCTATGCTGACAGA 1200
Db 1567 GGTGTGTGTGACCACTCAGCTGAGGCAAAACATCTATCCGATTCCTATGCTGACAGA 1626

```

```

QY 1201 TTCTTATGTTGAACAATTAGGAATGCTGGCCCAATCAAGGCAATTGATGTCGATACGG 1260
Db 1527 TTCTTATGTTGAACAATTAGGAATGCTGGCCCAATCAAGGCAATTGATGTCGATACGG 1686
QY 1261 CACTCATTTTGTGACCATGATACATCACTCTGTGACCAAGAAATTTTAAACCATTAACA 1320
Db 1687 CACTCATTTTGTGACCATGATACATCACTCTGTGACCAAGAAATTTTAAACCATTAACA 1746
QY 1321 GACTACTGCGCTCTCTCTCTCAAGCGCCTGATCGGCGCCCAAGTCTTGTGCTGAT 1380
Db 1747 GACTACTGCGCTCTCTCTCTCAAGCGCCTGATCGGCGCCCAAGTCTTGTGCTGAT 1806
QY 1381 GTGCTGCGGCTCCAGCGGAAACCAAGGCTGCGCGGAGTATCCGGAACAACTTAAGAT 1440
Db 1807 GTGCTGCGGCTCCAGCGGAAACCAAGGCTGCGCGGAGTATCCGGAACAACTTAAGAT 1866
QY 1441 TATGCTCAGTGCACAAACCAACACACACATAGTTGTGAGGTCATTAACATTTT 1500
Db 1867 TATGCTCAGTGCACAAACCAACACACACATAGTTGTGAGGTCATTAACATTTT 1926
QY 1501 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGAGTCTCAAGACAG 1560
Db 1927 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGAGTCTCAAGACAG 1986
QY 1561 CTGCTTACCAAGTACCTGCTGACGCGCTATGGGAGAGGCGCTTAAAGTCAAGTCACTG 1620
Db 1987 CTGCTTACCAAGTACCTGCTGACGCGCTATGGGAGAGGCGCTTAAAGTCAAGTCACTG 2046
QY 1621 CACTGATGCGCAGCCCTTATGATGATGAGACGACGCGGACCTCCCAAGATTAAGCCC 1680
Db 2047 CACTGATGCGCAGCCCTTATGATGATGAGACGACGCGGACCTCCCAAGATTAAGCCC 2106
QY 1681 CGCCCGCTTGGGCGCGGCGCGGACATTTGTCATCCCTGCAATCAATGGGCTTTTATGTG 1740
Db 2107 CGCCCGCTTGGGCGCGGCGCGGACATTTGTCATCCCTGCAATCAATGGGCTTTTATGTG 2166
QY 1741 GTCAAGATGTCATGCTTGTGGCTGCGGCTACCGATTA 1779
Db 2167 GTCAAGATGTCATGCTTGTGGCTGCGGCTACCGATTA 2205

RESULT 12
AAD29203
ID AAD29203 standard; cDNA; 1674 BP.
XX
AC AAD29203;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human heparanase-2 splice variant cDNA.
XX
KW Human; heparanase-2; Hep-2; wound healing; angiogenesis; restenosis;
KW atherosclerosis; neurodegenerative disease; inflammation; prostanine;
KW viral infection; autoimmune lesion; renal failure; pancreatic cancer;
KW dystrophic muscular disease; heart disease; gene therapy; enzyme; ss.
OS Homo sapiens.
XX
PN MO200204645-A2.
XX
PD 17-JAN-2002.
XX
PF 12-JUL-2001; 2001WO-EP008094.
XX
PR 12-JUL-2000; 2000EP-00202442.
XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI David G, Duer J;
XX
DR WPI, 2002-171719/22.
XX
PT Heparanase-2 polypeptides and polynucleotides, useful for useful in wound

```

healing, angiogenesis, and for treating restenosis, atherosclerosis, inflammation, neurodegenerative diseases, and viral infections.

Disclosure; Page 40-41; 54p; English.

The invention relates to human heparanase-2 (Hep-2) polypeptides and polynucleotides. Heparanase-2 protein is useful in wound healing, angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases, inflammation and viral infections, as well as in neutralising plasma heparin as a potential replacement of protamine. Antiheparanase-2 antibodies may be used for immunodetection and diagnosis of microvascular diseases, autoimmune lesions, renal failure in biopsy specimens, plasma samples and body fluids. Molecules which can agonise or antagonise heparanase-2 catalytic activity may also be used as a medicament. Polymorphisms in the polynucleotide sequence are useful in the identification of individuals having a predisposition to acquire diseases resulting from an increased or decreased expression of their activity. Such molecules can be used to treat pancreatic cancer, dystrophic muscular diseases and/or heart diseases. Polynucleotides of the invention are used in gene therapy. The present sequence is human heparanase-2 splice variant cDNA

Sequence 1674 BP; 419 A; 449 C; 421 G; 384 T; 0 U; 1 Other;

Query Match 79.3%; Score 1411; DB 6; Length 1674;

Best Local Similarity 89.8%; Pred. No. 0;

Matches 1598; Conservative 1; Mismatches 6; Indels 174; Gaps 1;

```

QY 1 ATGAGGAGTCTTTCGCTTCCCTGAGAGCCATGCTCCAGACATCCCGCCCGCCCGC 60
DB 56 ATGAGGAGTCTTTCGCTTCCCTGAGAGCCATGCTCCAGACATCCCGCCCGCCCGC 115
QY 61 TGCCTAGCCCGCGGCGCTCTACTTGGCTGTGCTCCATCTCTCTCTCTCTCCAG 120
DB 116 TGCCTAGCCCGCGGCGCTCTACTTGGCTGTGCTCCATCTCTCTCTCTCTCCAG 175
QY 121 GCTGAGACAGAGACCTTTCGCTGTAGACAGAGCTGCGAGTTTGAAGAAAAGCCTG 180
DB 176 GCTGAGACAGAGACCTTTCGCTGTAGACAGAGCTGCGAGTTTGAAGAAAAGCCTG 235
QY 181 ATTCTACTGATGTAGACCAAGAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 240
DB 236 ATTCTACTGATGTAGACCAAGAACCCAGTCAAGACAGTCAATGAGAACTTCTCTCT 295
QY 241 CTGCACTGATTCCTGATCATCATGATGCTGCTGATTTCTTAAAGCTTCAAGCGC 300
DB 296 CTGCACTGATTCCTGATCATCATGATGCTGCTGATTTCTTAAAGCTTCAAGCGC 355
QY 301 TTGGTACCTGCGCCGCGGGAATTTGCGCGGCTTCTGCGCTTGGGGGGCAAAAGAAC 360
DB 356 TTGGTACCTGCGCCGCGGGAATTTGCGCGGCTTCTGCGCTTGGGGGGCAAAAGAAC 415
QY 361 GACTTCTGCGAGTTCCAGAACCTGAGAACCCGCGGAAAAGCCGCGGCGCCCGC 420
DB 416 GACTTCTGCGAGTTCCAGAACCTGAGAACCCGCGGAAAAGCCGCGGCGCCCGC 475
QY 421 GATTACTATCTCAAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 476 GATTACTATCTCAAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 535
QY 481 CAGAAAAGCTGCAAGATGCCAGACCTGATGATGATGATGATGATGATGATGATGAT 540
DB 536 CAGAAAAGCTGCAAGATGCCAGACCTGATGATGATGATGATGATGATGATGATGAT 595
QY 541 GCGAGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 596 GCGAGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
QY 601 AATATTAAGCAAGGCTCTTGAACAACTTATTAATCTGATGATGATGATGATGAT 660
DB 656 AATATTAAGCAAGGCTCTTGAACAACTTATTAATCTGATGATGATGATGATGAT 664
QY 661 CTGATATTTTGTCTAATGACATGCGTGTATCCCAATTAATCTCGAAGAGTTCTAGT 720

```

```

DB 665 ----- 664
QY 721 GCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAAGTACATTTCTTGGAACTGGGT 780
DB 665 ----- 664
QY 781 AATGAGCCAAATACTATCTGAGACATGATGCGCGGCGATTAATGCGACCGATTGGA 840
DB 665 ---GAGCCAAATACTATCTGAGACATGATGCGCGGCGATTAATGCGACCGATTGGA 721
QY 841 AAGATTAATCACTGAGTGAAGAGCTGTGAGAGCCATCCGATTTATTCAGAGCCAGC 900
DB 722 AAGATTAATCACTGAGTGAAGAGCTGTGAGAGCCATCCGATTTATTCAGAGCCAGC 781
QY 901 TTATATGCGCTTAATATTTGGGCGCGCGAGAGAGATGATGATGATGATGATGATGAT 960
DB 782 TTATATGCGCTTAATATTTGGGCGCGCGAGAGAGATGATGATGATGATGATGATGAT 841
QY 961 ATGAGAGTGGCAGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 842 ATGAGAGTGGCAGAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGAT 901
QY 1021 CGGCTGATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 902 CGGCTGATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
QY 1081 ATTAGAAAATTCAGAAAAGTGTAAATACATCACTCCAGAAAAGAAATTTGGCTTGA 1140
DB 962 ATTAGAAAATTCAGAAAAGTGTAAATACATCACTCCAGAAAAGAAATTTGGCTTGA 1021
QY 1141 GGTGTGTGACCACTTCAAGTGAAGCAAAACATCTATCCGATTTCTATGCTGACGA 1200
DB 1022 GGTGTGTGACCACTTCAAGTGAAGCAAAACATCTATCCGATTTCTATGCTGACGA 1081
QY 1201 TTCTTATGTTGAACCTTTAGAAATGCTGCGCCATCAAGGCAATGATGCTGATACGG 1260
DB 1082 TTCTTATGTTGAACCTTTAGAAATGCTGCGCCATCAAGGCAATGATGCTGATACGG 1141
QY 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1142 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
QY 1321 GACTATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
DB 1202 GACTATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1261
QY 1381 GTGGCTGGGCTCGAGCGGAAGCCAGCGCTGCGCGAGTATCCGCGGACAACTTAAGATT 1440
DB 1262 GTGGCTGGGCTCGAGCGGAAGCCAGCGCTGCGCGAGTATCCGCGGACAACTTAAGATT 1321
QY 1441 TATGCTACATGCAAAACCAACCAACCAACCACTTATGCTGATGATGATGATGATGAT 1500
DB 1322 TATGCTACATGCAAAACCAACCAACCAACCACTTATGCTGATGATGATGATGATGAT 1381
QY 1501 ATCATCAACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1382 ATCATCAACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441
QY 1561 CTGGTTACCAAGTACTGCTGCAAGCCCTATGAGGACAGAGGCTTAAAGTCCAAAGTAA 1620
DB 1442 CTGGTTACCAAGTACTGCTGCAAGCCCTATGAGGACAGAGGCTTAAAGTCCAAAGTAA 1501
QY 1621 CAATGATGAGGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1502 CAATGATGAGGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
QY 1681 CGCCCTTTGGGCGCGCGGACATTTGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1562 CGCCCTTTGGGCGCGCGGACATTTGATGATGATGATGATGATGATGATGATGATGAT 1621
QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799

```

DB 1622 GTCAAGATGTCAATGCTTGGCTGCCGCTACCGATTAA 1660

RESULT 13  
AAH28347  
ID AAH28347 standard; cDNA; 3943 BP.  
AC AAH28347;  
XX 05-SEP-2001 (first entry)  
XX  
XX Nucleotide sequence of human heparanase-like polypeptide.  
DE  
XX Human; heparanase-like polypeptide; gene therapy; cancer; angiogenesis;  
XX trauma; autoimmune disease; skin disease; cardiovascular disease;  
XX nervous system disease; inflammation; arthritis; genitalia;  
XX male fertility; erectile dysfunction; ss.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1479  
XX /tag= a  
XX /transl\_except= (pos: 1219..1221, aa: Xaa)  
XX /note= "heparanase-like polypeptide; Xaa is an  
XX unspecified residue"  
XX  
XX MO200148161-A2.  
XX  
XX 05-JUL-2001.  
XX  
XX 18-DEC-2000; 2000MO-EP012909.  
XX  
XX 23-DEC-1999; 99EP-00125831.  
XX  
XX (SCHD ) SCHERING AG.  
XX  
XX Siemeister G, Weiss B;  
XX  
XX WPI: 2001-418259/44.  
XX  
XX P-PSDB: AAB84664.  
XX  
XX  
XX Human Heparanase-like polynucleotide encoding polypeptides useful for  
XX modulating expression of the polypeptide and for treating cancer, cancer  
XX metastasis, aberrant angiogenesis by gene therapy technique.  
XX  
XX Claim 1; Page 27-29; 30pp; English.  
XX  
XX The present sequence encodes a human heparanase-like polypeptide.  
XX Heparanase-like polynucleotides are useful as a source of probes, primers  
XX and antisense molecules, and in gene therapy. Heparanase-like  
XX polynucleotides and polypeptides are useful for treating several  
XX disorders e.g., cancer, cancer metastasis. The oligonucleotides are also  
XX useful as diagnostic markers for the diagnosis of disorder such as  
XX cancer, cancer metastasis and aberrant angiogenesis. They may also act as  
XX diagnostic markers for diagnosis of disorder such as cancer, cancer  
XX metastasis and aberrant angiogenesis. The heparanase polypeptides and  
XX polynucleotides are also useful for treating trauma, autoimmune diseases,  
XX skin diseases, cardiovascular diseases, nervous system diseases, and  
XX inflammation including arthritis. Since the polynucleotide is  
XX preferentially expressed in male genitalia, modulation of its expression  
XX and/or activity may be used for medical intervention in male genitalia  
XX function that is male fertility control, erectile dysfunction  
XX  
XX Sequence 3943 BP; 1089 A; 921 C; 996 G; 934 T; 0 U; 3 Other;  
SQ  
Query Match 72.5%; Score 1289.8; DB 5; Length 3943;  
Best Local Similarity 89.2%; Pred. No. 0;  
Matches 1475; Conservative 1; Mismatches 3; Indels 174; Gaps 1;  
QY 127 GACGAGAGACCTTGGCTGTGACAGAGCTGACAGTTTGAAGAAAAGACCTGATTCTA 186  
DB 1 GACGAGAGACCTTGGCTGTGACAGAGCTGACAGTTTGAAGAAAAGACCTGATTCTA 60

QY 187 CTTGATGTGAGACCAAGAAACCCAGTCAAGACAGTCAATGAAACTTCTCTCTGCGAG 246  
DB 61 CTTGATGTGAGACCAAGAAACCCAGTCAAGACAGTCAATGAAACTTCTCTCTGCGAG 120  
QY 247 CTGATCGCTGCATCATTCATGATGAGCTGAGCTGATTTCTTAAGCTCCAAAGCGTTGGTG 306  
DB 121 CTGATCGCTGCATCATTCATGATGAGCTGAGCTGATTTCTTAAGCTCCAAAGCGTTGGTG 180  
QY 307 ACCCTGCGCCG393AATTGCGCCGCTTCTGCGCTTGGG393CAAAAGACCGACTTC 366  
DB 181 ACCCTGCGCCG393AATTGCGCCGCTTCTGCGCTTGGG393CAAAAGACCGACTTC 240  
QY 367 CTGCAATTCAGAAACCTGAGAAACCCGCGAAACCCGCG393CCGCGCGCGGATTAC 426  
DB 241 CTGCAATTCAGAAACCTGAGAAACCCGCGAAACCCGCG393CCGCGCGGATTAC 300  
QY 427 TATCTCAAAAACCTATGAGATGATGATTTGCGAGTGTGCTTAGATPAAACAGAAA 486  
DB 301 TATCTCAAAAACCTATGAGATGATGATTTGCGAGTGTGCTTAGATPAAACAGAAA 360  
QY 487 GAGTCGAAGATTGCCAGACCCCTGATGTTATGCTGTGCTCCAAAGGAGAGAGAGCT 546  
DB 361 GAGTCGAAGATTGCCAGACCCCTGATGTTATGCTGTGCTCCAAAGGAGAGAGAGCT 420  
QY 547 CAGATGCAATGTGTTCTTCTTAAAGAGCAATTCCTCAATCTTACAGTAACTCATATTA 606  
DB 421 CAGATGCAATGTGTTCTTCTTAAAGAGCAATTCCTCAATCTTACAGTAACTCATATTA 480  
QY 607 ACAGCCAGGCTCTAGACAAAATTATTAACCTTGCTGATGCTCTGCACTCCACTGATA 666  
DB 481 ACA----- 483  
QY 667 TTGCTTAATGACATGCTGCTGTAATCCAAATACCTCTGGAACAGTTAGTCCCTG 726  
DB 484 ----- 483  
QY 727 AGTCTGTGAAGTACAGCGCAGCAAAAATTAACAACATTTCTTGGAACTGGGTAATGAG 786  
DB 484 -----GAG 486  
QY 787 CCAATTAATCTATCCGACCATGATGCGCGGAGTAATGAGCAGCAGTTGGAAAGAT 846  
DB 487 CCAATTAATCTATCCGACCATGATGCGCGGAGTAATGAGCAGCAGTTGGAAAGAT 546  
QY 847 TACATCCAGCTGAGAGAGCTGTTGCAAGCCCATCCGGAATTTATTCAGAGCCAGCTTATAT 906  
DB 547 TACATCCAGCTGAGAGAGCTGTTGCAAGCCCATCCGGAATTTATTCAGAGCCAGCTTATAT 606  
QY 907 GGCCTTAATTTGGGCGGCGCGAGAAAGATGTATCGCCCTCTAGATGATTCATGAG 966  
DB 607 GGCCTTAATTTGGGCGGCGCGAGAAAGATGTATCGCCCTCTAGATGATTCATGAG 666  
QY 967 GTGCGAGAGAGTACAGTATGATGAGTACTGCGCAACATTTCTCATTTGATGAGCGGAGTG 1026  
DB 667 GTGCGAGAGAGTACAGTATGATGAGTACTGCGCAACATTTCTCATTTGATGAGCGGAGTG 726  
QY 1027 GTCAAGGTATGAGATTTCTGAAAACCTGCTGTTAGACAACCTCTGACCAATTTAGG 1086  
DB 727 GTCAAGGTATGAGATTTCTGAAAACCTGCTGTTAGACAACCTCTGACCAATTTAGG 786  
QY 1087 AAAATTGGAAGGTGTTATATCATATACATCCAGGAAGAAAGATTTGGCTTGAAGGTGTG 1146  
DB 787 AAAATTGGAAGGTGTTATATCATATACATCCAGGAAGAAAGATTTGGCTTGAAGGTGTG 846  
QY 1147 GTGACCACTCAGCTGAGAGCAAAACATTCATTCGATGCTGACAGATTCTTA 1206  
DB 847 GTGACCACTCAGCTGAGAGCAAAACATTCATTCGATGCTGACAGATTCTTA 906  
QY 1207 TGGTTGAACACTTGAAGATGCTGCGCAATCAGGCAATGATGTGTATACGCGACTCA 1266  
DB 907 TGGTTGAACACTTGAAGATGCTGCGCAATCAGGCAATGATGTGTATACGCGACTCA 966



```

Db      534 ATCCGATTATTCACAGCCAGCTTATATGGCCCTAATTTGGGGGGCCGAGAGAGAAAT 593
Qy      937 GTGATGCCCTCTCTAGATGAGATTCATTAAGTGGCAGAGAAATGACATGATGCACTTACC 996
Db      594 GTGATGCCCTCTCTAGATGAGATTCATTAAGTGGCAGAGAAATGACATGATGCACTTACC 653
Qy      997 TGGCAACATTCATGATGATGAGCCGGGTGATCAAGGTGATGATGATCTCCGAAATCTGG 1056
Db      654 TGGCAACATTCATGATGATGAGCCGGGTGATCAAGGTGATGATGATCTCCGAAATCTGG 713
Qy      1057 CTGTTAGACACACTCTCTGACACAGATTAGAGAAATTCAGAAAGTGGTTAATACATCACT 1116
Db      714 CTGTTAGACACACTCTCTGACACAGATTAGAGAAATTCAGAAAGTGGTTAATACATCACT 773
Qy      1117 CCAGGAAGAAGATTGGCTTGAAGGTGGTGAACAACCTCAGCTGGAGGCAAAACAAT 1176
Db      774 CCAGGAAGAAGATTGGCTTGAAGGTGGTGAACAACCTCAGCTGGAGGCAAAACAAT 833
Qy      1177 CTATCCGATTCTATGCTGACAGATTCTTATGTTGAACAATTAGGAATGCTGGCCAAT 1236
Db      834 CTATCCGATTCTATGCTGACAGATTCTTATGTTGAACAATTAGGAATGCTGGCCAAT 893
Qy      1237 CAGGGCATTGATGCTGATACGGCACTCAATTTTGAACATGATGATCAATCACTCGTG 1296
Db      894 CAGGGCATTGATGCTGATACGGCACTCAATTTTGAACATGATGATCAATCACTCGTG 953
Qy      1297 GACCAAAATTTAAACCATTAACAGACATGCTGCTCTCTCTCTCAACAAGCCCTGATC 1356
Db      954 GACCAAAATTTAAACCATTAACAGACATGCTGCTCTCTCTCTCAACAAGCCCTGATC 1013
Qy      1357 GGGCCCAAACTTTGGTGTGCTGATGCTGGGCTCCAGGGAAGCAAGCCCTGGCCCA 1416
Db      1014 GGGCCCAAACTTTGGTGTGCTGATGCTGGGCTCCAGGGAAGCAAGCCCTGGCCCA 1073
Qy      1417 GTGATCCGGGACAAACTAAGATTTATGCTCACTGCAAAACCAACAACAACAACAACA 1476
Db      1074 GTGATCCGGGACAAACTAAGATTTATGCTCACTGCAAAACCAACAACAACAACAACA 1133
Qy      1477 GTTGTGGTGTGCTGATGATGCTTTTATCATCACTTGCATGATCAAGAAAGAAATCAAG 1536
Db      1134 GTTGTGGTGTGCTGATGATGCTTTTATCATCACTTGCATGATCAAGAAAGAAATCAAG 1193
Qy      1537 CTGGCTGGAGCTCTCAGAGACAAGCTGGTTCACAGATCTGCTGAGGCCCTATGGGCA 1596
Db      1194 CTGGCTGGAGCTCTCAGAGACAAGCTGGTTCACAGATCTGCTGAGGCCCTATGGGCA 1253
Qy      1597 GAGGGCTTAAGTCCAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1656
Db      1254 GAGGGCTTAAGTCCAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1313
Qy      1657 GGAACCTCCCAATTTGAAGCCCGCCCTTCGGGCGGGCGGACATTTGTCATCTCT 1716
Db      1314 GGAACCTCCCAATTTGAAGCCCGCCCTTCGGGCGGGCGGACATTTGTCATCTCT 1373
Qy      1717 CCAGTCAACATGAGGCTTTTATGCTCAAGATGATGATGATGATGATGATGATGATGAT 1776
Db      1374 CCAGTCAACATGAGGCTTTTATGCTCAAGATGATGATGATGATGATGATGATGAT 1433
Qy      1777 TAA 1779
Db      1434 TAA 1436

```

```

XX      Heparanase; hmbp1; wound healing; angiogenesis; restenosis; Scurvy;
KM      atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;
KM      neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;
KM      gene therapy; human; ds.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      CDS      26..1468
FT      /tag=a
FT      /product="heparanase"
PN      WO200100643-A2.
PD      04-JAN-2001.
PF      19-JUN-2000; 2000MO-IL000358.
PR      25-JUN-1999; 99US-0140801P.
PA      (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX      Pecker I, Michal I, Itzhaki H;
XX      WPI; 2001-137930/14.
DR      P-PSDB; AAY97634.
XX      New polynucleotides and polypeptides that are distantly homologous to
PT      heparanase, useful in wound healing, as well as in gene therapy protocols
PT      for angiogenesis, restenosis, atherosclerosis, or inflammation.
XX      Claim 3; Page 62; 67P; English.
XX      This sequence encodes a heparanase of the invention. The heparanase DNA
CC      and protein sequences are useful in wound healing, angiogenesis,
CC      restenosis, atherosclerosis, inflammation, pulmonary diseases,
CC      neurodegenerative diseases (such as Scurvy, Alzheimer's disease, and
CC      Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
CC      sequence is particularly useful in gene therapy.
XX      Sequence 1724 BP; 423 A; 481 C; 422 G; 398 T; 0 U; 0 Other;
SQ
Query Match      61.6%; Score 1095.4; DB 4; Length 1724;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 1; Indels 336; Gaps 1;
Qy      1 ATGAGGGTGTCTTTGTGCTTCCCTGAAGCCATGCGCTCCAGCACTCCGCCCCCGCG 60
Db      26 ATGAGGGTGTCTTTGTGCTTCCCTGAAGCCATGCGCTCCAGCACTCCGCCCCCGCG 85
Qy      61 TGGCTAGCCCCGGGGGCTCTTACTTGGCTGTGCTGCTCACTCTCCCTTCTCCAG 120
Db      86 TGGCTAGCCCCGGGGGCTCTTACTTGGCTGTGCTGCTCACTCTCCCTTCTCCAG 145
Qy      121 GCTGAGACAGAGAACCTTCCGTGTAGACAGAGCTCAGGTTTGAAGAAAGAACCTGT 180
Db      146 GCTGAGACAGAGAACCTTCCGTGTAGACAGAGCTCAGGTTTGAAGAAAGAACCTGT 205
Qy      181 ATTCTACTTGATGTGAGACCAAGAACCCAGTCAAGACGTCAATGAAATTCTCTCT 240
Db      206 ATTCTACTTGATGTGAGACCAAGAACCCAGTCAAGACGTCAATGAAATTCTCTCT 265
Qy      241 CTGACGCTGATTCGGTGCATCATGATGAGTGGCTGATTTCTTAAGTCCAAAGCG 300
Db      266 CTGACGCTGATTCGGTGCATCATGATGAGTGGCTGATTTCTTAAGTCCAAAGCG 325
Qy      301 TTGCTGACCTTGCGCCGGGGAATTGCGCCCGCTTTGCGCTTGGGGGCAAAAGAGCC 360
Db      326 TTGCTGACCTTGCGCCGGGGAATTGCGCCCGCTTTGCGCTTGGGGGCAAAAGAGCC 385
Qy      361 GACTTCCGAGCTTCCAGAACCTGAGAACCCGGCGAAAGCGCGGGGGCGCGGCGCG 420

```

RESULT 15  
 AAA91099  
 ID AAA91099 standard; DNA; 1724 BP.  
 AC AAA91099;  
 XX 20-APR-2001 (first entry)  
 DT  
 XX Human heparanase, hmbp1 pns form, coding sequence.





QY 181 ATTCTAATTGATGTAGACCAAGAAACCAAGTCAGACAGTCATAGAACTTCTCTCT 240  
 Db 181 ATTCTAATTGATGTAGACCAAGAAACCAAGTCAGACAGTCATAGAACTTCTCTCT 240  
 QY 241 CTGCACTGGATCCGTCATCATTCATGATGAGTGGCTGGATTTCTTAACTCAAGGCG 300  
 Db 241 CTGCACTGGATCCGTCATCATTCATGATGAGTGGCTGGATTTCTTAACTCAAGGCG 300  
 QY 301 TTGAGTACCTGAGCCCGGGGACCTTTCGCGCTTTCGCGCTTCGAGGAGCAAGGAC 360  
 Db 301 TTGAGTACCTGAGCCCGGGGACCTTTCGCGCTTTCGCGCTTCGAGGAGCAAGGAC 360  
 QY 361 GACTTCTGTAGTTCAGAACTGAGAAACCCGCGCAAAACCGCGGGGCGCGCGCGCG 420  
 Db 361 GACTTCTGTAGTTCAGAACTGAGAAACCCGCGCAAAACCGCGGGGCGCGCGCGCG 420  
 QY 421 GATTACTATCTCAAAACTATGAGATGACATTTGTCGAAGTGTGCTTCTTAAATAA 480  
 Db 421 GATTACTATCTCAAAACTATGAGATGACATTTGTCGAAGTGTGCTTCTTAAATAA 480  
 QY 481 CAGAAAGGCTGCAAGATTCGCCGACACCTGATGTTATGCTGCTCCAAAGGAGAG 540  
 Db 481 CAGAAAGGCTGCAAGATTCGCCGACACCTGATGTTATGCTGCTCCAAAGGAGAG 540  
 QY 541 GCAGCTCAGATGATCTGTTCTTCTTAAAGGAGCAATTCGCAATCTTACAGTAATCTC 600  
 Db 541 GCAGCTCAGATGATCTGTTCTTCTTAAAGGAGCAATTCGCAATCTTACAGTAATCTC 600  
 QY 601 ATATTAACAGCCAGGTCTCTAAGCAAACTTTATATCTTGTCTGATGCTGCACTCC 660  
 Db 601 ATATTAACAGCCAGGTCTCTAAGCAAACTTTATATCTTGTCTGATGCTGCACTCC 660  
 QY 661 CTGATATTTGCTCTAATGATGATGCGGTGATGCTGATGCTGATGCTGATGCTGAT 720  
 Db 661 CTGATATTTGCTCTAATGATGATGCGGTGATGCTGATGCTGATGCTGATGCTGAT 720  
 QY 721 GCCCTGAGTCTGTTGAAGTACAGGCGCAAGAAAGTACAACTTTCTTGGAACTGGGT 780  
 Db 721 GCCCTGAGTCTGTTGAAGTACAGGCGCAAGAAAGTACAACTTTCTTGGAACTGGGT 780  
 QY 781 AATGAGCCAAATACTATGAGACCATGATGATGCGCGGAGTAATAGGCAAGGCGGAG 840  
 Db 781 AATGAGCCAAATACTATGAGACCATGATGATGCGCGGAGTAATAGGCAAGGCGGAG 840  
 QY 841 AAGGATTAATCACTCAAGTGAAGAGCTGTTGCAAGCCATCCGATTTATTCAGAGCCAG 900  
 Db 841 AAGGATTAATCACTCAAGTGAAGAGCTGTTGCAAGCCATCCGATTTATTCAGAGCCAG 900  
 QY 901 TTAATAGCCCTAATATTTGGCGCGCGAGAGAAATGTCATGCGCTTCTCAATGAGATT 960  
 Db 901 TTAATAGCCCTAATATTTGGCGCGCGAGAGAAATGTCATGCGCTTCTCAATGAGATT 960  
 QY 961 ATGAAGGTGCAAGAACTATAGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 961 ATGAAGGTGCAAGAACTATAGATGATGATGATGATGATGATGATGATGATGATG 1020  
 QY 1021 CGGCTGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1080  
 Db 1021 CGGCTGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1080  
 QY 1081 ATTAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
 Db 1081 ATTAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
 QY 1141 GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 Db 1141 GGTGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 QY 1201 TTCTTATGTTGAACACTTTAGAAATGCTGGCAATCAAGGCAATGATGATGATGATG 1260  
 Db 1201 TTCTTATGTTGAACACTTTAGAAATGCTGGCAATCAAGGCAATGATGATGATGATG 1260

QY 1261 CACTATTTTTTGAACATGATATCAATCACTGCTGGAACAGATTTTAACCATTAACA 1320  
 Db 1261 CACTATTTTTTGAACATGATATCAATCACTGCTGGAACAGATTTTAACCATTAACA 1320  
 QY 1321 GACTACTGCTCTCTCTCTCTCTCAAGCGCTGATGCGGCCCAAGTCTTGGCTGAT 1380  
 Db 1321 GACTACTGCTCTCTCTCTCTCTCAAGCGCTGATGCGGCCCAAGTCTTGGCTGAT 1380  
 QY 1381 GTGCTGGGCTCCAGCGGAGACCAAGCGCTGCGCGGATGATCCGGAACAACATAGAT 1440  
 Db 1381 GTGCTGGGCTCCAGCGGAGACCAAGCGCTGCGCGGATGATCCGGAACAACATAGAT 1440  
 QY 1441 TATGCTACTGCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCACTTT 1500  
 Db 1441 TATGCTACTGCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCACTTT 1500  
 QY 1501 ATCATCAACTTGCATGATCAAGAAATAATCAAGTGGCTGGACTCTCAAGACAG 1560  
 Db 1501 ATCATCAACTTGCATGATCAAGAAATAATCAAGTGGCTGGACTCTCAAGACAG 1560  
 QY 1561 CTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
 Db 1561 CTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
 QY 1621 CACTGAATGCGACGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 Db 1621 CACTGAATGCGACGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 QY 1681 CGCCGCCCTGCGGCGCGGCGGAGATGATGATGATGATGATGATGATGATGATGAT 1740  
 Db 1681 CGCCGCCCTGCGGCGCGGCGGAGATGATGATGATGATGATGATGATGATGATGAT 1740  
 QY 1741 GTCAAGAAATGCAATGCTTTGGCTGCGCTACCGATTA 1779  
 Db 1741 GTCAAGAAATGCAATGCTTTGGCTGCGCTACCGATTA 1779

RESULT 2  
 US-10-177-245A-1  
 ; Sequence 1, Application US/10177245A  
 ; Publication No. US2003083254A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McKenzie, Edward Alexander  
 ; APPLICANT: Stamps, Alasdair Craig  
 ; APPLICANT: Terrell, Jonathan Alexander  
 ; APPLICANT: Tyson, Kerry Louise  
 ; TITLE OF INVENTION: Substances  
 ; FILE REFERENCE: 2543-1-027  
 ; CURRENT APPLICATION NUMBER: US/10/177,245A  
 ; PRIOR FILING DATE: 2002-08-29  
 ; PRIOR APPLICATION NUMBER: PCT/GB00/04963  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: UK 0008713.0  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: UK 9930392.7  
 ; PRIOR FILING DATE: 1999-12-22  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2636  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)..(2636)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-177-245A-1

Query Match 99.7%; Score 1773.8; DB 15; Length 2636;  
 Best Local Similarity 99.8%; Freq. No. 0;  
 Matches 1775; Conservative 4; Indels 0; Gaps 0;  
 QY 1 ATGAGGAGTGTGCTTGTGCTTGTGAGGATGCGCTTCAAGCAATCCCGCCCGCGG 60



D	b		1681	ATGTGAAATAATCAGAAAAGTGGTTAATATACATACACTCCAGGAAAAAGAATTGGCTTGA	1740
O	y		1141	GGTGTGTGACCACTCGACTGCTGAGGACAAGAACATCTATTCGCATTTCTCATCTCTCAGGA	1200
D	b		1741	GGTGGTGGACCACTCAGCTGAGGACAAACAAATCTATTCGCATTTCTCATCTCTCAGGA	1800
O	y		1201	TTCCTATGTTGAACACTTTAGGAATGCTGGCCATTCAGGGCATTGATTCGTGATACC	1260
D	b		1801	TTCCTATGTTGAACACTTTAGGAATGCTGGCCATTCAGGGCATTGATTCGTGATACC	1860
O	y		1261	CACCTATTTTTTGAACCATGATATACATCACTCTGTGACACAGAAITTTTAACCATTTACA	1320
D	b		1861	CACCTATTTTTTGAACCATGATATACATCACTCTGTGACACAGAAITTTTAACCATTTACA	1920
O	y		1321	GACTACTGGCTCTCTCTCTCTCTCTACAAAGGCCCTGATCCGGCCCCAAGTCTGGCTGTGAT	1380
D	b		1921	GACTACTGGCTCTCTCTCTCTCTCTACAAAGGCCCTGATCCGGCCCCAAGTCTGGCTGTGAT	1980
O	y		1381	GTGGCTGGGCTCCAGCGGAGCCACGGCCTGGCCGAGTAGTATCCGGACAAACTAGGATT	1440
D	b		1981	GTGGCTGGGCTCCAGCGGAGCCACGGCCTGGCCGAGTAGTATCCGGACAAACTAGGATT	2040
O	y		1441	TATGCTCACTGACAAACCAACCAACCACTCACTACGTTCTGGGGTCCATTAACATTTT	1500
D	b		2041	TATGCTCACTGACAAACCAACCAACCACTCACTACGTTCTGGGGTCCATTAACATTTT	2100
O	y		1501	ATCATCACTTGCATCGATCAAGAAAGAAATCAAAGCTGGTGGGACTCTCAGAGCAAG	1560
D	b		2101	ATCATCACTTGCATCGATCAAGAAAGAAATCAAAGCTGGTGGGACTCTCAGAGCAAG	2160
O	y		1561	CTGGTTCACTGACAGTACTGCTGTGACGCCCTATGGGAGGAGAGGGCCTAAGTCCAAGTAGTG	1620
D	b		2161	CTGGTTCACTGACAGTACTGCTGTGACGCCCTAAGTGGGAGGAGGGCCTAAGTCCAAGTAGTG	2220
O	y		1621	CAACTGAATGGCGAGCCCTTAGTATGTGTGACGAGCGGACCCCTCCAGAAITGAAGCCC	1680
D	b		2221	CAACTGAATGGCGAGCCCTTAGTATGTGTGACGAGCGGACCCCTCCAGAAITGAAGCCC	2280
O	y		1681	CGCCCCCTTGGGGCCGGCCGGAACATTGTGATCCTTCAGTCAACCATGGGCTTTTATGTG	1740
D	b		2281	CGCCCCCTTGGGGCCGGCCGGAACATTGTGATCCTTCAGTCAACCATGGGCTTTTATGTG	2340
O	y		1741	GTCAGAATGTCAATGCTTTGGGCTGCGGCTACCGATAA	1799
D	b		2341	GTCAGAATGTCAATGCTTTGGGCTGCGGCTACCGATAA	2379
 RESULT 3 US-10-302-172-409 Sequence 409, Application US/10302172 Publication No. US20040053250A1 GENERAL INFORMATION: APPLICANT: Tang, Y. Tom APPLICANT: Xue, Aifong J. APPLICANT: Dimaec, Radoje T. TITLE OF INVENTION: No. US004005250A1ei Arginine-rich Protein-like Nucleic Acids and TITLE OF INVENTION: Polypeptides FILE REFERENCE: 803_1CNCE CURRENT APPLICATION NUMBER: US/10/302,172 CURRENT FILING DATE: 2002-11-21 PRIOR APPLICATION NUMBER: US 10/225,251 PRIOR FILING DATE: 2002-08-20 PRIOR APPLICATION NUMBER: PCT US02/05095 PRIOR FILING DATE: 2002-03-05 PRIOR APPLICATION NUMBER: US 09/799,451 PRIOR FILING DATE: 2001-03-05 NUMBER OF SEQ ID NOS: 950 SOFTWARE: pc_FL_genes Version 2.0 SEQ ID NO 409 LENGTH: 1891 TYPE: DNA ORGANISM: Homo sapiens FEATURE:					

NAME/KEY: CDS  
LOCATION: (20)...(1621)  
US-10-302-172-409

Query Match 79.7%; Score 1417.8; DB 13; Length 1891;  
Best Local Similarity 90.1%; Pred. No. 0;  
Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

QY 1 ATAGAGGCTCTTGTGCTCTTCCCTGAGGCAATGCCCTCCAGCAATCCCGCCCGCCG 60  
DB 20 ATAGAGGCTCTTGTGCTCTTCCCTGAGGCAATGCCCTCCAGCAATCCCGCCCGCCG 79  
QY 61 TGGCTGAGGCTCTTGTGCTCTTCCCTGAGGCAATGCCCTCCAGCAATCCCGCCCGCCG 120  
DB 80 TGGCTGAGGCTCTTGTGCTCTTCCCTGAGGCAATGCCCTCCAGCAATCCCGCCCGCCG 139  
QY 121 GCTGAGAGCAGAGACCTTCTCTGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 140 GCTGAGAGCAGAGACCTTCTCTGAGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAG 199  
QY 181 ATTCTATCTGATGTGAGACCAAGAACCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 200 ATTCTATCTGATGTGAGACCAAGAACCAAGTCAAGAGAGAGAGAGAGAGAGAGAG 259  
QY 241 CTGAGCTGAGTCCGTCATCATTCATGATGAGTCTGATGATGATGATGATGATGATG 300  
DB 260 CTGAGCTGAGTCCGTCATCATTCATGATGAGTCTGATGATGATGATGATGATGATG 319  
QY 301 TTGCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAG 360  
DB 320 TTGCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAG 379  
QY 361 GACTCTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 420  
DB 380 GACTCTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 439  
QY 421 GATTACTATCTCAAAAATCTAGAGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 440 GATTACTATCTCAAAAATCTAGAGATGATGATGATGATGATGATGATGATGATGAT 499  
QY 481 CAGAAAGCTGCAAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
DB 500 CAGAAAGCTGCAAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559  
QY 541 GAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 560 GAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 619  
QY 601 ATTATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 620 ATTATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628  
QY 661 CTGATATTTCTCTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
DB 629 ----- 628  
QY 721 GCCCTGAGTCTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB 629 ----- 628  
QY 781 AATGAGCAATACTATGAG 840  
DB 629 ---GAGCAATACTATGAG 685  
QY 841 AAGGATTAATCCAGCTGAG 900  
DB 686 AAGGATTAATCCAGCTGAG 745  
QY 901 TTATATGAGCTATATTTGAG 960  
DB 746 TTATATGAGCTATATTTGAG 805  
QY 961 ATGAGAGTGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

DB 806 ATGAGAGTGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 865  
QY 1021 CGGCTGATCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
DB 866 CGGCTGATCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 925  
QY 1081 ATTGAGAAATTTGAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
DB 926 ATTGAGAAATTTGAGAAAGTGTATATATATATATATATATATATATATATATATAT 985  
QY 1141 GGTGTGTGACCACTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
DB 986 GGTGTGTGACCACTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045  
QY 1201 TTCTTATGTTGAAACATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1046 TTCTTATGTTGAAACATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1105  
QY 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 1106 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165  
QY 1321 GACTATGAGTCT 1380  
DB 1166 GACTATGAGTCT 1225  
QY 1381 GTGCTGAGGCTCCAGGAG 1440  
DB 1226 GTGCTGAGGCTCCAGGAG 1285  
QY 1441 TATGCTCATGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
DB 1286 TATGCTCATGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1345  
QY 1501 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
DB 1346 ATCATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1405  
QY 1561 CTGCTTCAACAGTACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1406 CTGCTTCAACAGTACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1465  
QY 1621 CACTGATGAG 1680  
DB 1466 CACTGATGAG 1525  
QY 1681 CGCCCTCTGAGGAG 1740  
DB 1526 CGCCCTCTGAGGAG 1585  
QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
DB 1586 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1624

RESULT 4  
US-09-836-461-1  
Sequence 1, Application US/09836461  
Patent No. US20020064853A1  
GENERAL INFORMATION:  
APPLICANT: Blenkowski, Michael J.  
APPLICANT: Heinrichson, Robert L.  
TITLE OF INVENTION: Heparanase II, A No. US20020064853A1 Human Heparanase Paralog  
FILE REFERENCE: heparanase II  
CURRENT APPLICATION NUMBER: US/09/836,461  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2326  
TYPE: DNA  
ORGANISM: Homo sapiens

US-09-836-461-1

Query Match 79.7%; Score 1417.8; DB 9; Length 2326;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

QY 1 ATGAGGCTGCTTTGCTGCTTCCCTGAAAGCATGCTCCCTCCAGCAACTCCCGCCCCCGCG 60  
 DB ATGAGGCTGCTTTGCTGCTTCCCTGAAAGCATGCTCCCTCCAGCAACTCCCGCCCCCGCG 84  
 QY 61 TGGCTAGCCCGGGGGGCTCTCACTTGGCTCTGTGCTCAATCTCCCTTCTCCAG 120  
 DB TGGCTAGCCCGGGGGGCTCTCACTTGGCTCTGTGCTCAATCTCCCTTCTCCAG 144  
 QY 121 GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGCAAGTTGAAGAAAAGACCTG 180  
 DB GCTGAGACAGAGACCTTGGCTGTAGACAGAGCTGCAAGTTGAAGAAAAGACCTG 204  
 QY 181 ATTCTAGTGTAGACACCAAGAACCCAGTCAAGTCAATGAGACTCTCTCT 240  
 DB ATTCTAGTGTAGACACCAAGAACCCAGTCAAGTCAATGAGACTCTCTCT 264  
 QY 241 CTGCACTGATCCGTCCATCATTCATGATGCTGCTGATTTCTTAAGTCCAGCC 300  
 DB CTGCACTGATCCGTCCATCATTCATGATGCTGCTGATTTCTTAAGTCCAGCC 324  
 QY 301 TTGGTACCTTGGCCCGGGGACTTTGGCCGCTTCTGGCTTCCGGGGGCAAAAGGAC 360  
 DB TTGGTACCTTGGCCCGGGGACTTTGGCCGCTTCTGGCTTCCGGGGGCAAAAGGAC 384  
 QY 325 TTGGTACCTTGGCCCGGGGACTTTGGCCGCTTCTGGCTTCCGGGGGCAAAAGGAC 420  
 DB TTGGTACCTTGGCCCGGGGACTTTGGCCGCTTCTGGCTTCCGGGGGCAAAAGGAC 444  
 QY 361 GACTCTCTGAGTTCCAGAACCTGAGGAAACCCGCGGAGGAGCCGCGGCGCG 480  
 DB GACTCTCTGAGTTCCAGAACCTGAGGAAACCCGCGGAGGAGCCGCGGCGCG 504  
 QY 421 GATTATCTATCAAAAATACTAGAGATGATTTGTAAGTATGTTGCTTAAGTAA 480  
 DB GATTATCTATCAAAAATACTAGAGATGATTTGTAAGTATGTTGCTTAAGTAA 504  
 QY 445 GATTATCTATCAAAAATACTAGAGATGATTTGTAAGTATGTTGCTTAAGTAA 540  
 DB GATTATCTATCAAAAATACTAGAGATGATTTGTAAGTATGTTGCTTAAGTAA 564  
 QY 481 CAGAAAGGCTGCAAGATTTGCCAGACCTGATGTTAAGTGTGCTCCAAAGGAGAA 540  
 DB CAGAAAGGCTGCAAGATTTGCCAGACCTGATGTTAAGTGTGCTCCAAAGGAGAA 564  
 QY 505 CAGAAAGGCTGCAAGATTTGCCAGACCTGATGTTAAGTGTGCTCCAAAGGAGAA 600  
 DB CAGAAAGGCTGCAAGATTTGCCAGACCTGATGTTAAGTGTGCTCCAAAGTCTC 624  
 QY 541 GCAAGCTCAATGATCTGCTTTCTTAAGAGAGAAATCTCCAAATCTTAAGTATCTC 624  
 DB GCAAGCTCAATGATCTGCTTTCTTAAGAGAGAAATCTCCAAATCTTAAGTATCTC 648  
 QY 601 AATATTACAGCAGGTCTTGAACAACCTTAATTAATTGCTGATGCTGTGACTCCAC 660  
 DB AATATTACAGCAGGTCTTGAACAACCTTAATTAATTGCTGATGCTGTGACTCCAC 684  
 QY 625 AATATTACAGCAGGTCTTGAACAACCTTAATTAATTGCTGATGCTGTGACTCCAC 720  
 DB AATATTACAGCAGGTCTTGAACAACCTTAATTAATTGCTGATGCTGTGACTCCAC 744  
 QY 661 CTGATATTTGCTTAAATGCACTGCTGATTCCTCAATACTCTGGAACAGTTCTAGT 780  
 DB CTGATATTTGCTTAAATGCACTGCTGATTCCTCAATACTCTGGAACAGTTCTAGT 804  
 QY 721 GCGCTAGCTCTGTGAAGTACAGCGCCAGCAAAAAGTACAACATTTCTGGAACTGGGT 840  
 DB GCGCTAGCTCTGTGAAGTACAGCGCCAGCAAAAAGTACAACATTTCTGGAACTGGGT 864  
 QY 781 AATGAGCCAAATACTATCGGACCATGATGCTGCGGCACTTAATGCGAGCAGTTGGGA 840  
 DB AATGAGCCAAATACTATCGGACCATGATGCTGCGGCACTTAATGCGAGCAGTTGGGA 864  
 QY 841 AAGGATTAATCCAGCTGAGAGAGCTGTGCAAGCCCATCCGGATTTATTCAGAGCCAGC 900  
 DB AAGGATTAATCCAGCTGAGAGAGCTGTGCAAGCCCATCCGGATTTATTCAGAGCCAGC 924  
 QY 901 TTATATGCGCTATATTTGGCGCGCCAGAGAAAGATGCAATCGCCCTCTAAGTGAATTC 960  
 DB TTATATGCGCTATATTTGGCGCGCCAGAGAAAGATGCAATCGCCCTCTAAGTGAATTC 984  
 QY 961 ATGAAGGTGAGAGAGATGATGATGCACTTCTGGCAACTTGTCTATGATGAGC 1020  
 DB ATGAAGGTGAGAGAGATGATGATGCACTTCTGGCAACTTGTCTATGATGAGC 1044  
 QY 811 ATGAAGGTGAGAGAGATGATGATGCACTTCTGGCAACTTGTCTATGATGAGC 870

QY 1021 CCGGTGCTCAAGTGTATGATCTTCTGAAAACCTCGCTTGAACAACCTCTGACCGAG 1080  
 DB CCGGTGCTCAAGTGTATGATCTTCTGAAAACCTCGCTTGAACAACCTCTGACCGAG 930  
 QY 1081 ATTGAGAAAATTGAGAAAGTGTATATACATACCTCCAGAAAAGAAATTTGGCTTGA 1140  
 DB ATTGAGAAAATTGAGAAAGTGTATATACATACCTCCAGAAAAGAAATTTGGCTTGA 990  
 QY 931 ATTGAGAAAATTGAGAAAGTGTATATACATACCTCCAGAAAAGAAATTTGGCTTGA 990  
 DB ATTGAGAAAATTGAGAAAGTGTATATACATACCTCCAGAAAAGAAATTTGGCTTGA 990  
 QY 1141 GGTGTGTGACCACTTGAAGTGTAGAAATGCTGGCCCAATGAGGAAATGATGATAGC 1200  
 DB GGTGTGTGACCACTTGAAGTGTAGAAATGCTGGCCCAATGAGGAAATGATGATAGC 1050  
 QY 1201 TTCTATGTTGAAACCTTGAAGTGTAGAAATGCTGGCCCAATGAGGAAATGATGATAGC 1260  
 DB TTCTATGTTGAAACCTTGAAGTGTAGAAATGCTGGCCCAATGAGGAAATGATGATAGC 1110  
 QY 1051 TTCTATGTTGAAACCTTGAAGTGTAGAAATGCTGGCCCAATGAGGAAATGATGATAGC 1110  
 DB TTCTATGTTGAAACCTTGAAGTGTAGAAATGCTGGCCCAATGAGGAAATGATGATAGC 1260  
 QY 1261 CACTCATTTTGTGACATGATATCAATCACTCTGTGACAGAAATTTTAAACCATTAACA 1320  
 DB CACTCATTTTGTGACATGATATCAATCACTCTGTGACAGAAATTTTAAACCATTAACA 1170  
 QY 1111 CACTCATTTTGTGACATGATATCAATCACTCTGTGACAGAAATTTTAAACCATTAACA 1170  
 DB CACTCATTTTGTGACATGATATCAATCACTCTGTGACAGAAATTTTAAACCATTAACA 1320  
 QY 1321 GACTATGCTGCTCTCTCTCTCTTCTTACAGAGCTGATGCGCCCAAAAGTGGCTGAT 1380  
 DB GACTATGCTGCTCTCTCTCTCTTCTTACAGAGCTGATGCGCCCAAAAGTGGCTGAT 1230  
 QY 1171 GACTATGCTGCTCTCTCTCTCTTCTTACAGAGCTGATGCGCCCAAAAGTGGCTGAT 1230  
 DB GACTATGCTGCTCTCTCTCTCTTCTTACAGAGCTGATGCGCCCAAAAGTGGCTGAT 1380  
 QY 1381 GTGGCTGGGCTCCAGCGGAAACCAAGGCTGCGGAGTATCCGGAACAACTAAGGAT 1440  
 DB GTGGCTGGGCTCCAGCGGAAACCAAGGCTGCGGAGTATCCGGAACAACTAAGGAT 1290  
 QY 1231 GTGGCTGGGCTCCAGCGGAAACCAAGGCTGCGGAGTATCCGGAACAACTAAGGAT 1290  
 DB GTGGCTGGGCTCCAGCGGAAACCAAGGCTGCGGAGTATCCGGAACAACTAAGGAT 1440  
 QY 1441 TATGCTGCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 DB TATGCTGCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1350  
 QY 1291 TATGCTGCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1350  
 DB TATGCTGCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500  
 QY 1501 ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 DB ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1410  
 QY 1351 ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1410  
 DB ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 QY 1561 CTGGTTCACAGTACCTGCTGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620  
 DB CTGGTTCACAGTACCTGCTGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1470  
 QY 1411 CTGGTTCACAGTACCTGCTGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1470  
 DB CTGGTTCACAGTACCTGCTGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620  
 QY 1621 CAAGTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB CAAGTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1530  
 QY 1471 CAAGTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1530  
 DB CAAGTGAATGAGCAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1680  
 QY 1681 CGCCCTTGGGCGGCGCGGAGCAATGATGATGATGATGATGATGATGATGATGATG 1740  
 DB CGCCCTTGGGCGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATG 1590  
 QY 1531 CGCCCTTGGGCGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATG 1590  
 DB CGCCCTTGGGCGGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATG 1740  
 QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1799  
 DB GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1629

## RESULT 5

US-10-177-245A-3

Sequence 3, Application US/10177245A

Publication No. US20030083254A1

GENERAL INFORMATION:

APPLICANT: McKenzie, Edward Alexander

APPLICANT: Stamps, Alasdair Craig

APPLICANT: Terrett, Jonathan Alexander

APPLICANT: Tyson, Kerry Louise

TITLE OF INVENTION: Substances

FILE REFERENCE: 2543-1-027

CURRENT APPLICATION NUMBER: US/10/177,245A

CURRENT FILING DATE: 2002-08-29

PRIOR APPLICATION NUMBER: PCT/GB00/04963

PRIOR FILING DATE: 2000-12-21

PRIOR APPLICATION NUMBER: UK 0008713.0

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: UK 9930392.7

PRIOR FILING DATE: 1999-12-22

NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 3  
 LENGTH: 2462  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (1)-(2462)  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-177-245A-3

Query Match 79.7%; Score 1417.8; DB 15; Length 2462;

Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1603; Conservative 0; Mismatches 2; Indels 174; Gaps 1;

1 ATGAGGGTGTCTTGTGCTTCCCTTCCCTGAAAGCATGCTCCGAGCACTCCCGCCCGCCCGCG 60  
 601 ATGAGGGTGTCTTGTGCTTCCCTTCCCTGAAAGCATGCTCCGAGCACTCCCGCCCGCGCG 660  
 61 TGCCTAGCCCGCGGGGCTCTCTACTTGTGCTGCTCCATCTCTCCCTTCTCTCCAG 120  
 661 TGCCTAGCCCGCGGGGCTCTCTACTTGTGCTGCTCCATCTCTCCCTTCTCTCCAG 720  
 121 GCTGAGACAGAGACCTTCTGCTGAGACAGAGCTGAGGTTTGAAGAAAAGCCTG 180  
 721 GCTGAGACAGAGACCTTCTGCTGAGACAGAGCTGAGGTTTGAAGAAAAGCCTG 780  
 181 ATTCTACTGATGAGACCAAGAACCAAGTCAAGACAGTCAATGAACTTCTCT 240  
 781 ATTCTACTGATGAGACCAAGAACCAAGTCAAGACAGTCAATGAACTTCTCT 840  
 241 CTGACCTGATCGTCCATCATTCATGATGCTGCTGCTGCTTCTTAAAGTCCAGGC 300  
 841 CTGACCTGATCGTCCATCATTCATGATGCTGCTGCTGCTTCTTAAAGTCCAGGC 900  
 301 TTGGTACCTGCGCGGGGACCTTTCGCGCGCTTCTGCGCTGCGGGGCAAAAGGAC 360  
 901 TTGGTACCTGCGCGGGGACCTTTCGCGCGCTTCTGCGCTGCGGGGCAAAAGGAC 960  
 361 GACTTCCTGAGTTCAGAACCTGAGAACCCCGGCAAAAGCCGCGGGGCGCGCGCG 420  
 961 GACTTCCTGAGTTCAGAACCTGAGAACCCCGGCAAAAGCCGCGGGGCGCGCGCG 1020  
 421 GATTACTATCTCAAAACTATGAGATGATGCTTGAAGTGTGCTTGAATAA 480  
 1021 GATTACTATCTCAAAACTATGAGATGATGCTTGAAGTGTGCTTGAATAA 1080  
 481 CAGAAAGCTGCAAGATGCTCCAGCACTGATGTTATGCTGCTCCAAAGGAGAG 540  
 1081 CAGAAAGCTGCAAGATGCTCCAGCACTGATGTTATGCTGAGCTCCAAAGGAGAG 1140  
 541 GCGAGCTCAGATGATCTGCTTCTTAAAGAGAGCATTTCTCAATCTTACAGTATCTC 600  
 1141 GCGAGCTCAGATGATCTGCTTCTTAAAGAGAGCATTTCTCAATCTTACAGTATCTC 1200  
 601 ATATTACAGCAGCTCTAGACAACTTATTACTTGTGATGCTCTGAGCTCCAG 660  
 1201 ATATTACAGCAGCTCTAGACAACTTATTACTTGTGATGCTCTGAGCTCCAG 1209  
 661 CTGATATTGCTCTAAATGACCTGCGTGAATCCCAATTACTCTGAGAACGTTCTAGT 720  
 1210 ----- 1209  
 721 GCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAGTAAACATTTCTGGAGACTGGGT 780  
 1210 ----- 1209  
 781 AATGAGCAATAATCTATGAGACCATGATGCGCGGAGTAAATGAGCAGCCAGTTGGGA 840  
 1210 -----GAGCAATAATCTATGAGACCATGATGCGCGGAGTAAATGAGCAGCCAGTTGGGA 1266  
 841 AAGGATTACATCAGCTGAGAGAGCTGTGACAGCCCATCGGATTAATTCAGAGCCAGC 900

1267 AAGGATTACATCAGCTGAGAGAGCTGTGACAGCCCATCGGATTAATTCAGAGCCAGC 1326  
 901 TTAATAGCCCTTAATATTGCGCGCGCGAGAAAGATGATGCGCCCTCTAGATGATTC 960  
 1327 TTAATAGCCCTTAATATTGCGCGCGCGAGAAAGATGATGCGCCCTCTAGATGATTC 1386  
 961 ATGAGGTGACAGAAATACAGTATGAGATGATGATGATGATGATGATGATGATGATG 1020  
 1387 ATGAGGTGACAGAAATACAGTATGAGATGATGATGATGATGATGATGATGATGATG 1446  
 1021 CCGGTGATCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 1447 CCGGTGATCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1506  
 1081 ATTAGAAATTCAGAAAGTGTAAATATACATACATCCAGGAAAGATTTGGCTTGA 1140  
 1507 ATTAGAAATTCAGAAAGTGTAAATATACATACATCCAGGAAAGATTTGGCTTGA 1566  
 1141 GGTGATGACACCTCAGCTGAGAGACCAACATCTATCCGATTCCTATGCTGAGGA 1200  
 1567 GGTGATGACACCTCAGCTGAGAGACCAACATCTATCCGATTCCTATGCTGAGGA 1626  
 1201 TTTCTATGTTGAACCTTAGGATGCTGAGGACCAATGAGGACATGATGCTGATACGG 1260  
 1627 TTTCTATGTTGAACCTTAGGATGCTGAGGACCAATGAGGACATGATGCTGATACGG 1686  
 1261 CACTATTTTGTGACATGAGATCAATCACCTGAGGACCAATTTTAAACCATTAACA 1320  
 1687 CACTATTTTGTGACATGAGATCAATCACCTGAGGACCAATTTTAAACCATTAACA 1746  
 1321 GACTACTGCT 1380  
 1747 GACTACTGCT 1806  
 1381 GTGCTGAGGCTGACAGCGGAAGCAAGCGCTGAGGATGAGTCCGGAACAACTAAGATT 1440  
 1807 GTGCTGAGGCTGACAGCGGAAGCAAGCGCTGAGGATGAGTCCGGAACAACTAAGATT 1866  
 1441 TATGCTACTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1500  
 1867 TATGCTACTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1926  
 1501 ATCATCACTTGCATCATCAAGAAAGAAATCAAGCTGAGGACCTCTCAGAGCAAG 1560  
 1927 ATCATCACTTGCATCATCAAGAAAGAAATCAAGCTGAGGACCTCTCAGAGCAAG 1986  
 1561 CTGCTTACCAATGCTGCTGACAGCTTATGAGGACAGAGGCTTAAAGTCAAGTCAAG 1620  
 1987 CTGCTTACCAATGCTGCTGACAGCTTATGAGGACAGAGGCTTAAAGTCAAGTCAAG 2046  
 1621 CACTGATAGGCGAGCCCTTAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
 2047 CACTGATAGGCGAGCCCTTAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2106  
 1681 CGCCCTCTGAGGCGCGGCGGCAATGATGATGATGATGATGATGATGATGATGATG 1740  
 2107 CGCCCTCTGAGGCGCGGCGGCAATGATGATGATGATGATGATGATGATGATGATG 2166  
 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 2167 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2205

RESULT 6  
 US-09-880-262-1

; Sequence 1, Application US/09880262  
 ; Patent No. US20020137907A1

; GENERAL INFORMATION:  
 ; APPLICANT: SOUTHAN, CHRISTOPHER DONALD  
 ; APPLICANT: LARMINIE, CHRISTOPHER GEOFFREY CARSON  
 ; APPLICANT: RANCE, KIM  
 ; APPLICANT: HAYES, PHILIP DAVID

; TITLE OF INVENTION: NOVEL COMPOUNDS  
 ; FILE REFERENCE: GP-30231  
 ; CURRENT APPLICATION NUMBER: US/09/880,262  
 ; CURRENT FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: UK 0014447.7  
 ; PRIOR FILING DATE: 2000-06-13  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO: 1  
 ; LENGTH: 1602  
 ; TYPE: DNA  
 ; ORGANISM: HOMO SAPIENS  
 ; US-09-880-262-1

Query Match 79.6%; Score 1416.4; DB 9; Length 1602;  
 Best Local Similarity 90.1%; Pred. No. 0;  
 Matches 1601; Conservative 0; Mismatches 1; Indels 174; Gaps 1;

QY 1 ATGAGGGTGTCTGTGCTTCCCTGAAAGCCATGCGCCCTCCAGCAACTCCGCCCCCGCG 60  
 DB 1 ATGAGGGTGTCTGTGCTTCCCTGAAAGCCATGCGCCCTCCAGCAACTCCGCCCCCGCG 60  
 QY 61 TGCCTAGCCCGGCGGCTCTCTAATTGCTCTGCTGCTCTCTCTCTCTCTCTCTCT 120  
 DB 61 TGCCTAGCCCGGCGGCTCTCTAATTGCTCTGCTCTCTCTCTCTCTCTCTCTCT 120  
 QY 121 GCTGAGACAGAGAACCTTGGCTGTGACAGAGCTGAGGCTTGAAGGAAAGACCTG 180  
 DB 121 GCTGAGACAGAGAACCTTGGCTGTGACAGAGCTGAGGCTTGAAGGAAAGACCTG 180  
 QY 181 ATTCTATTGATGTGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCTCT 240  
 DB 181 ATTCTATTGATGTGAGACCAAGAACCCAGTCAGACAGTCATGAGAACTTCTCTCT 240  
 QY 241 CTGCACTGATCCCTGATCATTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 DB 241 CTGCACTGATCCCTGATCATTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 301 TTGGTGAACCTGCGCGGGAATTTGGCGGCTTTCTGCGCTTCCGAGGCAAAAGAGC 360  
 DB 301 TTGGTGAACCTGCGCGGGAATTTGGCGGCTTTCTGCGCTTCCGAGGCAAAAGAGC 360  
 QY 361 GACTTCTGCAATTCGAAACCTGAGAACCCGCGGAAAGCCGCGGCGCGCGCGCG 420  
 DB 361 GACTTCTGCAATTCGAAACCTGAGAACCCGCGGAAAGCCGCGGCGCGCGCGCG 420  
 QY 421 GATTACTATCTCAAAAACCTATGAGATGACATTTGCTGAGATGATGCTTGAATAA 480  
 DB 421 GATTACTATCTCAAAAACCTATGAGATGACATTTGCTGAGATGATGCTTGAATAA 480  
 QY 481 CAGAAAGCTGCAAGATTTGCCAGCACCTGATGTTATGCTGAGCTCCAAAGGAGAG 540  
 DB 481 CAGAAAGCTGCAAGATTTGCCAGCACCTGATGTTATGCTGAGCTCCAAAGGAGAG 540  
 QY 541 GCACTCAGATGATCTGCTGCTTCTTAAAGAGCAATTCCTCAATCTTCAATATCTC 600  
 DB 541 GCACTCAGATGATCTGCTGCTTCTTAAAGAGCAATTCCTCAATCTTCAATATCTC 600  
 QY 601 AATATTAACAGCCAGGCTCTGAGCAAACTTTATTAATTGCTGATGCTGCTCCAC 660  
 DB 601 AATATTAACAGCCAGGCTCTGAGCAAACTTTATTAATTGCTGATGCTGCTCCAC 660  
 QY 661 CTGATATTGCTCTAAATGCACTGCGCTGTAATCCCAATAACTCTGGAACAGTTCTAGT 720  
 DB 661 CTGATATTGCTCTAAATGCACTGCGCTGTAATCCCAATAACTCTGGAACAGTTCTAGT 720  
 QY 721 GCCCTGAGTGTGTAAGTACAGGCGCAGCAAAAGTAAACATTTCTTGGAACTGGGT 780  
 DB 721 GCCCTGAGTGTGTAAGTACAGGCGCAGCAAAAGTAAACATTTCTTGGAACTGGGT 780  
 QY 781 AATGAGCAATAAATACTATGAGACATGCAATGCGCGGCAATAATGCAAGCACTTGGGA 840  
 DB 781 AATGAGCAATAAATACTATGAGACATGCAATGCGCGGCAATAATGCAAGCACTTGGGA 840  
 QY 840 ---GAGCAATAAATACTATGAGACATGCAATGCGCGGCAATAATGCAAGCACTTGGGA 866  
 DB 840 ---GAGCAATAAATACTATGAGACATGCAATGCGCGGCAATAATGCAAGCACTTGGGA 866

QY 841 AAGATTATCATCCAGCTGAGAGAGCTTGTGACAGCCCATCCGATTTATCCAGAGCCAGC 900  
 DB 841 AAGATTATCATCCAGCTGAGAGAGCTTGTGACAGCCCATCCGATTTATCCAGAGCCAGC 900  
 QY 901 TTATATGAGCCCTAATATTGGGCGGCGGAGAGAAATGTCAATGCGCTCTAGATGATTC 960  
 DB 901 TTATATGAGCCCTAATATTGGGCGGCGGAGAGAAATGTCAATGCGCTCTAGATGATTC 960  
 QY 961 ATGAGGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB 961 ATGAGGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 QY 1021 CGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 DB 1021 CGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 QY 1081 ATGAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATATAT 1140  
 DB 1081 ATGAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATATAT 1140  
 QY 1141 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
 DB 1141 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
 QY 1201 TTCTTATGTTGAACACTTTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 DB 1201 TTCTTATGTTGAACACTTTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 QY 1261 CACTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 DB 1261 CACTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 QY 1321 GACTACTGCT 1380  
 DB 1321 GACTACTGCT 1380  
 QY 1381 GTGGCTGGGCTCCAGCGGAGCCAGCGCTGGCGGAGTATCCGGAACAACTTAAGATT 1440  
 DB 1381 GTGGCTGGGCTCCAGCGGAGCCAGCGCTGGCGGAGTATCCGGAACAACTTAAGATT 1440  
 QY 1441 TATGCTCATGCAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1500  
 DB 1441 TATGCTCATGCAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1500  
 QY 1501 ATCATCACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 DB 1501 ATCATCACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 QY 1561 CTGCTTCAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 DB 1561 CTGCTTCAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 QY 1621 CACTGATGAG 1680  
 DB 1621 CACTGATGAG 1680  
 QY 1681 CGCCCTTGGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
 DB 1681 CGCCCTTGGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
 QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 DB 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 QY 1801 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
 DB 1801 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860

RESULT 7  
 ; US-09-978-249-2  
 ; Sequence 2, Application US/09978249  
 ; Patent No. US20020106780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PISCALIA, et al.  
 ; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies



APPLICANT: Stamps, Alasdair Craig  
 APPLICANT: Terrett, Jonathan Alexander  
 APPLICANT: Tyson, Kerry Louise  
 TITLE OF INVENTION: Substances  
 FILE REFERENCE: 2543-1-027  
 CURRENT APPLICATION NUMBER: US/10/177,245A  
 CURRENT FILING DATE: 2002-08-29  
 PRIOR APPLICATION NUMBER: PCT/GB00/04963  
 PRIOR FILING DATE: 2000-12-21  
 PRIOR APPLICATION NUMBER: UK 0008713.0  
 PRIOR FILING DATE: 2000-04-07  
 PRIOR APPLICATION NUMBER: UK 9930392.7  
 PRIOR FILING DATE: 1999-12-22  
 NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5  
 LENGTH: 2300  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(2300)  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-177-245A-5

Query Match 61.6%; Score 1095.4; DB 15; Length 2300;  
 Best Local Similarity 81.1%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

QY 1 ATGAGGAGTGTGTTGCTTGCCTTCCCTGAGCCATGCGCTCCGACACTCCGCCCCCGGG 60  
 DB 601 ATGAGGAGTGTGTTGCTTGCCTTCCCTGAGCCATGCGCTCCGACACTCCGCCCCCGGG 660  
 QY 61 TGCTTACCCCGGGGGCTCTCTACTTGGCTCTGTGCTCATCTTCCCTTTCTCCAG 120  
 DB 661 TGCTTACCCCGGGGGCTCTCTACTTGGCTCTGTGCTCATCTTCCCTTTCTCCAG 720  
 QY 121 GCTGAGACAGAGACCCCTTGCCTGTAGACAGAGCTGAGGTTGAAGAAAGACCTG 180  
 DB 721 GCTGAGACAGAGACCCCTTGCCTGTAGACAGAGCTGAGGTTGAAGAAAGACCTG 780  
 QY 181 ATTCTACTGATGTGACACCAAGAACCCAGTCAAGACATCAATGAACTTCTCTCT 240  
 DB 781 ATTCTACTGATGTGACACCAAGAACCCAGTCAAGACATCAATGAACTTCTCTCT 840  
 QY 241 CTGAGCTGATCCGTCCATCATTCATGATGCGTGGTGAATTCCTAAGCTCCAGCG 300  
 DB 841 CTGAGCTGATCCGTCCATCATTCATGATGCGTGGTGAATTCCTAAGCTCCAGCG 900  
 QY 301 TTGGTGAACCTTGGCCCGGGAGCTTTCCGCCCTTTCTGCGCTTCGCGGGCAAAAGAC 360  
 DB 901 TTGGTGAACCTTGGCCCGGGAGCTTTCCGCCCTTTCTGCGCTTCGCGGGCAAAAGAC 960  
 QY 361 GACTCTCTGAGTTCGAGAACCTGAGAACCCGCGGAAAGCCGCGGGGGCCGCGG 420  
 DB 961 GACTCTCTGAGTTCGAGAACCTGAGAACCCGCGGAAAGCCGCGGGGGCCGCGG 1020  
 QY 421 GATTACTATCTCAAAAACCTAGAGATGATTTGGAAGTATGTTGCTTAGATAA 480  
 DB 1021 GATTACTATCTCAAAAACCTAGAGATGATTTGGAAGTATGTTGCTTAGATAA 1045  
 QY 481 CAGAAAGCTGCAAGATTTGCCAGACACCTGATGTTATGCTGATGCTCCAAAGGAGAG 540  
 DB 1046 ----- 1045  
 QY 541 GCAAGCTCAGATGATCTGTTCTTTCTAAAGAGCAATTCCTCAATCTTACAGTAATCTC 600  
 DB 1046 ----- 1045  
 QY 601 ATATTACAGCCAGGCTCTAGACAAACTTTATACTTGGTGAATGCTCTGACCTCCAC 660  
 DB 1046 ----- 1045

QY 661 CTGATATTGCTCTPAATGCACTGCGTGPATCCCAATTACTCTGGAAACAGTTCTAGT 720  
 DB 1046 ----- 1045  
 QY 721 GCCCTGAGTCTGTGAAGTACAGGCCAGCAAAAAGTACAACATTTCTTGGGAATGGGT 780  
 DB 1046 ----- 1045  
 QY 781 AATGAGCCAAATTAATACTGAGCAGCATGAGCCGCGAGTAATGACAGCCAGTTGGGA 840  
 DB 1046 -ATGAGCCAAATTAATACTGAGCAGCATGAGCCGCGAGTAATGACAGCCAGTTGGGA 1104  
 QY 841 AAGATTATCATCAGCTGAAGAAGCTTGTGAGCCCATCCGGAATTTTCCAGAGCCAGC 900  
 DB 1105 AAGATTATCATCAGCTGAAGAAGCTTGTGAGCCCATCCGGAATTTTCCAGAGCCAGC 1164  
 QY 901 TTAATAGCCCTTAATATTGAGCGGCGCGAGAGAAATGTCAATCGCCCTCTAGATGATTC 960  
 DB 1165 TTAATAGCCCTTAATATTGAGCGGCGCGAGAGAAATGTCAATCGCCCTCTAGATGATTC 1224  
 QY 961 ATGAAAGTGGCAGAGATGACATGATGCACTTACTGGCAACATTTGCTACATTTAGTGC 1020  
 DB 1225 ATGAAAGTGGCAGAGATGACATGATGCACTTACTGGCAACATTTGCTACATTTAGTGC 1284  
 QY 1021 CGGATGATCAAGTGTAGTGAATCTTCTGAAAACCTGCGCTGTAGACACACTCTGACAG 1080  
 DB 1285 CGGATGATCAAGTGTAGTGAATCTTCTGAAAACCTGCGCTGTAGACACACTCTGACAG 1344  
 QY 1081 ATTAGAAAATTCAGAAAATGCTTATATCATCACTCCAGAAAAGAAATTTGGCTTGA 1140  
 DB 1345 ATTAGAAAATTCAGAAAATGCTTATATCATCACTCCAGAAAAGAAATTTGGCTTGA 1404  
 QY 1141 GGTGTGTGACCACTGACGCTGAGAGGACCAAAACATCTATCGATTCCTATGCTGACGA 1200  
 DB 1405 GGTGTGTGACCACTGACGCTGAGAGGACCAAAACATCTATCGATTCCTATGCTGACGA 1464  
 QY 1201 TTCTTATGTTGAACAATTGAGAAATGCTGGCAATCAAGGCAATGATGTGATACGG 1260  
 DB 1465 TTCTTATGTTGAACAATTGAGAAATGCTGGCAATCAAGGCAATGATGTGATACGG 1524  
 QY 1261 CACTGATTTTGGACATGATTAACAATCACTCGTGAACGAAATTTTAACCACTTAACA 1320  
 DB 1525 CACTGATTTTGGACATGATTAACAATCACTCGTGAACGAAATTTTAACCACTTAACA 1384  
 QY 1321 GACTACTGCTCTCTCTCTCTTACAGCGCCTGATCGGCCCAAAAGTCTTGGCTGAT 1380  
 DB 1585 GACTACTGCTCTCTCTCTCTCTTACAGCGCCTGATCGGCCCAAAAGTCTTGGCTGAT 1644  
 QY 1381 GTGGCTGGGCTCCAGCGGAGACCAAGGCTGCGGAGTATGCTGGGACCAACTAAGAT 1440  
 DB 1645 GTGGCTGGGCTCCAGCGGAGACCAAGGCTGCGGAGTATGCTGGGACCAACTAAGAT 1704  
 QY 1441 TATGCTCACTGACAAAACCAACAACAACAAGTATGCTGGGACCAACTAAGAT 1500  
 DB 1705 TATGCTCACTGACAAAACCAACAACAACAAGTATGCTGGGACCAACTAAGAT 1764  
 QY 1501 ATCATCACTGATGATCAAGAAAGAAATCAAGCTGCTGGGACCTCAAGACCAAG 1560  
 DB 1765 ATCATCACTGATGATCAAGAAAGAAATCAAGCTGCTGGGACCTCAAGACCAAG 1824  
 QY 1561 CTGATTCACAGTATCTGCTGAGCGCTATGAGGAGAGAGGCTTAAGCTCAAGTCAAGT 1620  
 DB 1825 CTGATTCACAGTATCTGCTGAGCGCTATGAGGAGAGAGGCTTAAGCTCAAGTCAAGT 1884  
 QY 1621 CACTGAATGAGCAGCCCTTATGATGATGAGAGAGAGGCTTCAAGTCAAGTCAAGT 1680  
 DB 1885 CACTGAATGAGCAGCCCTTATGATGATGAGAGAGAGGCTTCAAGTCAAGTCAAGT 1944  
 QY 1681 CGCCCGCTTGGGCGGCGGACATTTGATCAATCCCTCAAGTCAAGTCAAGTCAAGT 1740  
 DB 1945 CGCCCGCTTGGGCGGCGGACATTTGATCAATCCCTCAAGTCAAGTCAAGTCAAGT 2004  
 QY 1741 GTCAAGAAATGTCATGCTTTGGCTGCGCTACCAATTA 1779



Db 2005 GTCAAGATGTCATGCTTTGGCTGGCTGCGCTACCGATTA 2043

RESULT 9

US-10-080-254-53  
 ; Sequence 53, Application US/10080254  
 ; Publication No. US20030199008A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PA114C1  
 ; CURRENT APPLICATION NUMBER: US/10/080,254  
 ; CURRENT FILING DATE: 2002-02-22  
 ; NUMBER OF SEQ ID NOS: 169  
 ; Prior Application removed - See File Wrapper or Palm  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 53  
 ; LENGTH: 1114  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-080-254-53

Query Match 35.5%; Score 632.4; DB 15; Length 1114;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-199;  
 Matches 633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1146 GGTACACCTGACCTGAGGACAAACATCTATCCGATCTGACGAGATCTT 1205  
 Db 12 GGTACACCTGACCTGAGGACAAACATCTATCCGATCTGACGAGATCTT 71  
 Qy 1206 ATGGTTGAACCTTTAGGAATGCTGGCCAAATCAGGGCATTGATGCTGTAGTACGCACTC 1265  
 Db 72 ATGGTTGAACCTTTAGGAATGCTGGCCAAATCAGGGCATTGATGCTGTAGTACGCACTC 131  
 Qy 1266 ATTTTGGACATGATATCACTCTGCGACCAATTTTAACTTACCAACTA 1325  
 Db 132 ATTTTGGACATGATATCACTCTGCGACCAATTTTAACTTACCAACTA 191  
 Qy 1326 CTGGCTCTCTCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 1385  
 Db 192 CTGGCTCTCTCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 251  
 Qy 1386 TGGGCTCCAGGGAAGCCAGGCTGGCCGAGTATCCGGAACAATAAGATTATGC 1445  
 Db 252 TGGGCTCCAGGGAAGCCAGGCTGGCCGAGTATCCGGAACAATAAGATTATGC 311  
 Qy 1446 TCACTGACAAACCAACCAACAACAATAGTTGGTTCATTAACCTTTTATCAT 1505  
 Db 312 TCACTGACAAACCAACCAACAACAATAGTTGGTTCATTAACCTTTTATCAT 371  
 Qy 1506 CAATTGATGATCAAGAAAATAATCAAGTGTGCTGGAATCTCTCAGAGACAAGTGGT 1565  
 Db 372 CAATTGATGATCAAGAAAATAATCAAGTGTGCTGGAATCTCTCAGAGACAAGTGGT 431  
 Qy 1566 TCACTGATCTGCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 1385  
 Db 432 TCACTGATCTGCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 251  
 Qy 1626 GAATGCGAAGCCTTATGATGTGAGCAAGCGGACCTTCCAGAAATTGAAGCCCGCC 1685  
 Db 492 GAATGCGAAGCCTTATGATGTGAGCAAGCGGACCTTCCAGAAATTGAAGCCCGCC 551  
 Qy 1686 CCTTGGGCGGAGCGGACATTTGGCTCATCCCTCCAGTCAACCATGAGGCTTTTATGAGTCAA 1745  
 Db 552 CCTTGGGCGGAGCGGACATTTGGCTCATCCCTCCAGTCAACCATGAGGCTTTTATGAGTCAA 611  
 Qy 1746 GAATGTCATGCTTTGGCTGCGGATACCGATTA 1779  
 Db 612 GAATGTCATGCTTTGGCTGCGGATACCGATTA 645

RESULT 10

US-10-242-355-322  
 ; Sequence 322, Application US/10242355  
 ; Publication No. US20030235831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC003C1  
 ; CURRENT APPLICATION NUMBER: US/10/242,355  
 ; CURRENT FILING DATE: 2002-09-13  
 ; PRIOR APPLICATION NUMBER: 09/764,897  
 ; PRIOR FILING DATE: 2001-01-17  
 ; PRIOR APPLICATION NUMBER: 60/179,065  
 ; PRIOR FILING DATE: 2000-01-31  
 ; PRIOR APPLICATION NUMBER: 60/180,628  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: 60/214,886  
 ; PRIOR FILING DATE: 2000-06-28  
 ; PRIOR APPLICATION NUMBER: 60/217,487  
 ; PRIOR FILING DATE: 2000-07-11  
 ; PRIOR APPLICATION NUMBER: 60/225,758  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/220,963  
 ; PRIOR FILING DATE: 2000-07-26  
 ; PRIOR APPLICATION NUMBER: 60/217,496  
 ; PRIOR FILING DATE: 2000-07-11  
 ; PRIOR APPLICATION NUMBER: 60/225,447  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/218,290  
 ; PRIOR FILING DATE: 2000-07-14  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1267  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 322  
 ; LENGTH: 1114  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-242-355-322

Query Match 35.5%; Score 632.4; DB 16; Length 1114;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-199;  
 Matches 633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1146 GGTACACCTGACCTGAGGACAAACATCTATCCGATCTGACGAGATCTT 1205  
 Db 12 GGTACACCTGACCTGAGGACAAACATCTATCCGATCTGACGAGATCTT 71  
 Qy 1206 ATGGTTGAACCTTTAGGAATGCTGGCCAAATCAGGGCATTGATGCTGTAGTACGCACTC 1265  
 Db 72 ATGGTTGAACCTTTAGGAATGCTGGCCAAATCAGGGCATTGATGCTGTAGTACGCACTC 131  
 Qy 1266 ATTTTGGACATGATATCACTCTGCGACCAATTTTAACTTACCAACTA 1325  
 Db 132 ATTTTGGACATGATATCACTCTGCGACCAATTTTAACTTACCAACTA 191  
 Qy 1326 CTGGCTCTCTCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 1385  
 Db 192 CTGGCTCTCTCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 251  
 Qy 1386 TGGGCTCCAGGGAAGCCAGGCTGGCCGAGTATCCGGAACAATAAGATTATGC 1445  
 Db 252 TGGGCTCCAGGGAAGCCAGGCTGGCCGAGTATCCGGAACAATAAGATTATGC 311  
 Qy 1446 TCACTGACAAACCAACCAACAACAATAGTTGGTTCATTAACCTTTTATCAT 1505  
 Db 312 TCACTGACAAACCAACCAACAACAATAGTTGGTTCATTAACCTTTTATCAT 371  
 Qy 1506 CAATTGATGATCAAGAAAATAATCAAGTGTGCTGGAATCTCTCAGAGACAAGTGGT 1565  
 Db 372 CAATTGATGATCAAGAAAATAATCAAGTGTGCTGGAATCTCTCAGAGACAAGTGGT 431  
 Qy 1566 TCACTGATCTGCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 1385  
 Db 432 TCACTGATCTGCTCTCTCAAGCCCTGATCGGCCCAAGTCTGGCTGTGATGCG 251



QY 1626 GAATGCGAGCCCTTAGTGTGAGACGACGGGACCCCTCCAGATTGAAGCCCGCC 1685  
DB 492 GAATGCGAGCCCTTAGTGTGAGACGACGGGACCCCTCCAGATTGAAGCCCGCC 551  
QY 1686 CCTTCGGGCGCGCGGACATTTGATCCCTCCAGTACCATGAGGCTTTTATGTGTCAA 1745  
DB 552 CCTTCGGGCGCGCGGACATTTGATCCCTCCAGTACCATGAGGCTTTTATGTGTCAA 611  
QY 1746 GAATGTCAATGCTTTGGCTTGGCCCTTACCGATTA 1779  
DB 612 GAATGTCAATGCTTTGGCTTGGCCCTTACCGATTA 645

RESULT 11  
US-10-177-245A-22  
Sequence 22, Application US/10177245A  
Publication No. US20030083254A1  
GENERAL INFORMATION:  
APPLICANT: McKenzie, Edward Alexander  
APPLICANT: Stamps, Alasdair Craig  
APPLICANT: Terrett, Jonathan Alexander  
APPLICANT: Tyson, Kerry Louise  
TITLE OF INVENTION: Substances  
FILE REFERENCE: 2543-1-027  
CURRENT APPLICATION NUMBER: US/10/177,245A  
CURRENT FILING DATE: 2002-08-29  
PRIOR APPLICATION NUMBER: PCT/GB00/04963  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: UK 0008713.0  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: UK 9930392.7  
PRIOR FILING DATE: 1999-12-22  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 400  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(400)  
OTHER INFORMATION: n = A,T,C or G  
US-10-177-245A-22

Query Match 21.1%; Score 375.8; DB 15; Length 400;  
Best Local Similarity 96.8%; Pred. No. 8.5e-114;  
Matches 396; Conservative 0; Mismatches 4; Indels 9; Gaps 1;

QY 427 TATCTCAAAAACCTATGAGATGATGATGTTGAGATGATGTTGCTTAAATTAACAGAAA 486  
DB 1 TATCTCAAAAACCTATGAGATGATGATGTTGAGATGATGTTGCTTAAATTAACAGAAA 60  
QY 487 GCGTGCAGATTTGCCAGCAACCTGATGTTATGCTGTCTCCAAAGAGAGAGAGAGCT 546  
DB 61 GCGTGCAGATTTGCCAGCAACCTGATGTTATGCTGTCTCCAAAGAGAGAGAGAGCT 120  
QY 547 CAGATGATCTGTTCTTCTTAAAGAGCAATTCCTCAATCTTAACATTAATTCATTTA 606  
DB 121 CAGATGATCTGTTCTTCTTAAAGAGCAATTCCTCAATCTTAACATTAATTCATTTA 180  
QY 607 ACAGCCAGGCTCTAGACAAACTTATTAACCTTTGCTGATGCTTGAATCCACTGATA 666  
DB 181 ACAGCCAGGCTCTAGACAAACTTATTAACCTTTGCTGATGCTTGAATCCACTGATA 240  
QY 667 TTTCTCTAATGACGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 726  
DB 241 TTTCTCTAATGACGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 300  
QY 727 AGTCTGTGAAGTACAGGCGGAGCAAAAGTAAACAATTTCTTGGGAAGCTGGGTATGAG 786  
DB 301 AGTCTGTGAAGTACAGGCGGAGCAAAAGTAAACAATTTCTTGGGAAGCTGGGT----- 354

QY 787 CCAATTAATCTATCGAGCATGATGCGCGGAGCAATTAATGACAGCAGT 835  
DB 355 ---AATACTATCGAGCATGATGCGCGGAGCAATTAATGACAGCAGT 400

RESULT 12  
US-09-978-249-16  
Sequence 16, Application US/09978249  
Patent No. US20020106780A1  
GENERAL INFORMATION:  
APPLICANT: Fischeila, et al.  
TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies  
FILE REFERENCE: PTO54PI  
CURRENT APPLICATION NUMBER: US/09/978,249  
CURRENT FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: PCT/US01/11643  
PRIOR FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 60/198,123  
PRIOR FILING DATE: 2000-04-18  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 336  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-978-249-16

Query Match 18.8%; Score 334.4; DB 9; Length 336;  
Best Local Similarity 99.7%; Pred. No. 4.7e-100;  
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 448 GACATTTGTCGAGTGTGCTTAAATTAACAGAAAGCTGCAATTTGCCAGAC 507  
DB 1 GACATTTGTCGAGTGTGCTTAAATTAACAGAAAGCTGCAATTTGCCAGAC 60  
QY 508 CCTGATGTATGCTGTGCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 567  
DB 61 CCTGATGTATGCTGTGCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120  
QY 568 AAGAGCAATTTCTCCAAATCTTACAGTATCTGATTTAAGCAGGCTCTTACAGAA 627  
DB 121 AAGAGCAATTTCTCCAAATCTTACAGTATCTGATTTAAGCAGGCTCTTACAGAA 180  
QY 628 CTTTAACTTGTGATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 687  
DB 181 CTTTAACTTGTGATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 240  
QY 688 CGTAATCCCAATTAATCTCTGAGACAGTCTAGTGCCTGAGTCTGTAAGTACAGCGC 747  
DB 241 CGTAATCCCAATTAATCTCTGAGACAGTCTAGTGCCTGAGTCTGTAAGTACAGCGC 300  
QY 748 AGCAAAAAGTACACATTTCTTGGGAAGTGGTAT 783  
DB 301 AGCAAAAAGTACACATTTCTTGGGAAGTGGTAT 336

RESULT 13  
US-09-930-218-10  
Sequence 10, Application US/09930218  
Patent No. US20020034810A1  
GENERAL INFORMATION:  
APPLICANT: Goldsmith, Orit  
APPLICANT: Decker, Iris  
APPLICANT: Vlodevsky, Israel  
APPLICANT: Israel, Michael  
TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H  
FILE REFERENCE: 01/2235  
CURRENT APPLICATION NUMBER: US/09/930,218  
CURRENT FILING DATE: 2001-08-16  
PRIOR APPLICATION NUMBER: 09/666,390  
PRIOR FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 16



```

QY 920 GGCGGCCGAGGAAGATGTCATGCCCTCTAGATGATTCAGAGGTCGACGAGAGA 979
DB 768 GGCGAGCCCCGAGGACACACCCAGCAGCTCTCAGAACTTCAGAAATTCGAGGGAAG 827
QY 980 CAGTAGATGACATTCCTGCGCAACATTCCTACATTCATGATGCGCGGTCGTCAGTATGG 1039
DB 828 CCAATGACCTCGGTCACCTGCGCAACACTACTATGATGAAATGCGGAGTGCACGAGGAGG 887
QY 1040 ACTTCTGAAATGCGCTGTTGACACACTCTCTGACCAAGATTCGAGAAATTCAGAAAG 1099
DB 888 ATTTCCTGAGCCCTGAAAGTCGTGAGTCTCTTGGCACTGCCATACAGATTCCTGGGGA 947
QY 1100 TGCTTAATCATATCATCTCAGAGAAAGATTCGCTTGAAGGTCGTGACCACTCAG 1159
DB 948 TCCTGGAAGGAAAGGTCGCCGCGAGAGAGTATGCTGGGTGAGACCGGCTCGGCTACG 1007
QY 1160 CTGAGGCGCAAAACATCTATCCGATTCTTATGCTGCAAGATTCCTTATGTTGAACATT 1219
DB 1008 GCGGCGGCGCGCCCGCAGCTCTCCACACCTATGTCGCGGCTTCATGTCGTCGACCAAGC 1067
QY 1220 TAGGAAATGCTGGCCATCAGGCGATTGATGTCGTATACGCGCACTCATTTTTCACATG 1279
DB 1068 TGCGGTTGCGGCTCGGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1127
QY 1280 GATACATACACTCTGTCGACGAGAAATTTTAACTTACCACTACCTACCTCTCTCTCC 1339
DB 1128 GCACTATACCTCTGTCGATGTCGCGCTTCAAGCCCTTCCGCGACTCTGTCGTCATCTGC 1187
QY 1340 TCTACAAAGGCTGATGTCGCGCCCAAGTCTGTCGTCGATGTCGTCGTCGTCGTCGTCG 1399
DB 1188 TATACAAAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1247
QY 1400 AGCCAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1459
DB 1248 GCGGCGCGCG-----GGGTCATCTGCACTGCAACCAAC 1280
QY 1460 ACCACACCACTACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1519
DB 1281 CCGGCGACCCCAATACCGGGAAGGAGTGTACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1340
QY 1520 CAGAAAGAAATCAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1579
DB 1341 TGACCCAGACTTGCAGCTCTCTTAAGCAGTTCGAGTAAAGTTCGAGTAAAGTTCGAGTAAAG 1400
QY 1580 TCGAGCTCTTTCGCGAGGAGGCTTAAAGTTCGAGTAAAGTTCGAGTAAAGTTCGAGTAAAG 1639
DB 1401 TCGTCCCCCAGCGGACAGGACAGCATCTGTCAGAGAGTGCAGTGAATGCGCGCTTAC 1460
QY 1640 TAGGATGTCGAGGACGAGGAGCCCTCCAGATTCGAGGAGCCCGCTTCGCGCGCGCG 1699
DB 1461 TCGAGATGTCGAGGACGAGGACGAGTCCCGCGCTGTCAGAGAGTGCAGTGAATGCGCGCTTAC 1520
QY 1700 GAGATTCGTCATCTCTCAGTACCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1759
DB 1521 GCAAGCTCGCGCTGCGAGCTCTCTTACGCTTTCATGATGATCAGGAAAGCTTAAAGCTTA 1580
QY 1760 TGGCGTGC 1767
DB 1581 TTGCTTGC 1588

```

```

RESULT 15
US-09-776-874A-43
; Sequence 43, Application US/09776874A
; Patent No. US20020102560A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Iris
; APPLICANT: Vlodevsky, Israel
; APPLICANT: Feinseid, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 01/22603
; CURRENT APPLICATION NUMBER: US/09/776, 874A

```

```

; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 08/922,170
; PRIOR FILING DATE: 1997-09-02
; PRIOR APPLICATION NUMBER: US 09/109,386
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: PCT/US98/17954
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 2396
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-776-874A-43

Query Match 15.9%; Score 282.6; DB 9; Length 2396;
Best Local Similarity 52.8%; Pred. No. 3.2e-82;
Matches 677; Conservative 0; Mismatches 574; Indels 30; Gaps 2;

QY 474 AGATPAAACAGAAAGCTGCAAGATTGCCGACGACCTGATGTTATGTCGTGCTCCAAAG 533
DB 932 AGTCAACCATGATATTTGACAGTCTGACCGGTCCTGCTGCGGTGTCGAAATCCCA 991
QY 534 GGAGAGGCGAGCTCAGATGATCGTGTCTTCTTAAAGAGCAATTCATCTTCAATCTTACAG 593
DB 992 GGTGAAATGCGCTTCCAGAGCTGTTGCTGCTGCTCCAGAGCAGTACCAAAAGAGTTCA 1051
QY 594 TATCTCATATTAACAGCCAGGTCCTAGACAACTTATPACTTGTGATTTGCTCTCG 653
DB 1052 GAAACAGCACTACTCAAGAAAGCTCAGTGACATGCTCTACAGTTTTCGAAAGTCTCGG 1111
QY 654 ACTCAGCTGATTTTGTCTTAAATGCAAGCGCTGCTTAAATCCCAATACCTCGGAGACAG 713
DB 1112 GTTGAACCTGATCTTGTGTTAAATGCTTACAGACCCAGACTTACGATGAGGACAG 1171
QY 714 TTCTAGTCCCTGAGTCTGTTGAATGACAGCGCCAGCAAAAGTACAACTTTCTTGGGA 773
DB 1172 CTCACAGCCAGCTTCTCTGATCTGCTGCTTCAAGGGTTATPAACTTCTCTGGGA 1231
QY 774 ACTGGTATGAGCCAAATPACTTTCGACATGCGCGGCACTTAAATGCGACCA 833
DB 1232 ACTGGCAGTACGCGCAACAGTTTCTGAAAGAACTTCACTTCTCATGATGAGTTGCA 1291
QY 834 GTTGGAAAGATTAATCACTCAGTGAAGGCTGTGACGCCATCCGATTTATTCAG 893
DB 1292 GTTGAAGAAATCTTTGTGAGTTGCAATTAATCTTCA--AAGTCAGTTTCCAAA 1348
QY 894 AGCCAGTTTATGAGCTTAAATTTGCGCGCGCGAGGAAATGTCATCGCCCTCTTGA 953
DB 1349 TGCAAAATCTTATGCTCTGACATGTCAGCTCGAGGGAAGCAGTTAACTGCTGAG 1408
QY 954 TGGAATTCAAGAGGTCGAGAGATGACATGATGACATTCCTGCGCACTTCTCAT 1013
DB 1409 GAGTTTCTGAAAGGTCGCGAGAGATGATGATCTCTTACATGAGTACATCTTACTT 1468
QY 1014 TGATGCGCGGTCGTCAGAGTATGATCTTCAAAAATCTGCGCTTTAGACACATCTC 1073
DB 1469 GAATGAGCAGCTCGCTACCAAGAAATTTCTAGCTCTGATGCGCTGAGCACCTTTAT 1528
QY 1074 TGACAGATTCAGAAATTCAGAAAGTTCATATACATCACTCAGAGAAAGAAATTTG 1133
DB 1529 TCTCTCTGCAAAATTTCTGAAAGTCACTTAAAGATTCACACTGCGCAAGAGTCTG 1588
QY 1134 GCTTGAAGTGTGTCACCTCAGCTGAGGCGCAAAACATCTATTCGATTCCTATGC 1193
DB 1589 GTTGAAGAGCAGACTCAGCTTACGTCGTCGAGCTTGTCTGTCACACCTTTGC 1648
QY 1194 TGCAAGATTCATGCTTGAACATCTTGAAGATGTCGCAATCGGCACTGATGTCGT 1253
DB 1649 ACTGAGCTTATGTCGTCGATTAATTTGGCCTGTCAGCCAGATGAGGATGAGTGT 1708
QY 1254 GATAGGCACTCATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1313

```

Fri May 14 13:01:54 2004

us-10-088-676-1.rnpb

Page 14

Db	1709	GATAGGCAAGTGTCTTCCGAGCAGGCACTACCACTTAAGTATGAAACTTTGAGCC	1768
Qy	1314	ATTACCAAGACTACTGGCTCTCTCTCTCTACTAAGAGCCTGATTCGAGCCCAAAAGCTTGGC	1373
Db	1769	TTTATCCGTATTACTGGCTCTCTCTCTCTCTGTTCAAGAAACTGGTAGGTCACAGGGTGTACT	1828
Qy	1374	TGTGCACTGTGGCTGGGCTCCAGCGGAGGCCACGGGCTGGCCGAGTGATCCGGGCAAACT	1433
Db	1829	GTCAAGAGTGAAAGGCCAGA-----CAGAGCAAACT	1861
Qy	1434	AAGATTTATGCTCACTGCAACAACCAACAACAACCAACTAGTGGTGCATTAC	1493
Db	1862	CCGAGTGATGTCCACTGCACTTAAGTGTTACCCACGAATACAGAAAGAGATCTAAC	1921
Qy	1494	ACTTTTATCATCAACTGCATGCATCAAGAAAGAAATCAAGCTGGAGCTCTCAG	1553
Db	1922	TCTGTATCTCTGAACCTCCATTAATGTACCAAGACTTGAAGTACGGCTCCGTTGTT	1981
Qy	1554	AGACAACTGTGTTACCAAGTACTGCTGACAGCCCTATGGGCAAGAGGCTCTAAAGTCCA	1613
Db	1982	CAGGAAACAGAGTGATAGTACTTCTTAAAGCTTCGGGCGCGAGTGGATTACTTCCAA	2041
Qy	1614	GTCAGTGCACGTGAATGGCCAGCCCTTAGTAGTGTGACAGACGGGACCTCTCCAGAAAT	1673
Db	2042	ATCTGTCCAACTGAAGCGTCAATTCTGAAGTGTGATGACGAAACCTGCCACGCTT	2101
Qy	1674	GAAAGCCCCGCCCTTGGGGCCGAGCAGTGTGATCCCTCCAGTCAACCATGGGCTT	1733
Db	2102	GACAGAAAAACCTCTCCCCGAGGAAGTGCATAAGCCTGCCTGCTTTTCTCTATGCTTT	2161
Qy	1734	TTATGTGTCAGAAATGTCAA	1754
Db	2162	TTTTTCTATGAAGAAATGCCAA	2182

Search completed: May 13, 2004, 17:38:08  
Job time : 796 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 11:30:45 ; Search time 4726 Seconds  
(without alignments)  
11240.969 Million cell updates/sec

Title: US-10-088-676-1

Perfect score: 1779  
Sequence: 1 atgagagtgcttgcctt.....tggcctgcgcaccgataa 1779

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hlc: \*  
9: gb\_estl: \*  
10: gb\_est2: \*  
11: gb\_hlc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vtl: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rtd: \*  
26: em\_gss\_phg: \*  
27: em\_gss\_vtl: \*  
28: gb\_gss1: \*  
29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	703.6	39.6	784	14	CF726388
2	505.4	28.4	628	10	BB116041
3	420.2	23.6	623	12	BM426194
4	387.4	21.8	620	10	BB627984

Result No.	Score	Query Match	Length	DB ID	Description
5	306	17.2	573	12	BM083129
6	284.2	16.0	2173	11	AK040471
7	274.4	15.4	640	29	CE365946
8	258.8	14.5	351	9	AI019269
9	242	13.6	1962	11	AK087283
10	241.2	13.0	321	12	BI661209
11	232.6	12.4	888	13	BU474796
12	202.2	11.4	489	29	CG611539
13	191.6	10.8	1158	29	AL552151
14	187.6	10.5	334	29	CG653785
15	170.8	9.6	1200	9	AL545232
16	170.4	9.6	348	13	BX280673
17	166.6	9.4	427	29	AG019565
18	166.6	9.4	427	29	AG019565
19	162.4	9.1	777	9	AL718139
20	160	9.0	629	14	CB483444
21	150.6	8.5	631	14	CF915055
22	148.6	8.4	708	13	BQ75819
23	148.6	8.4	826	14	CK017201
24	144.6	8.1	433	9	AL917968
25	144.4	8.1	1083	13	BX398409
26	143.6	8.1	914	13	CG436333
27	139.6	7.8	285	29	CG486333
28	137.8	7.7	576	12	BQ037157
29	135.6	7.6	739	13	BU452690
30	130.2	7.3	907	13	BQ438834
31	126	7.1	652	13	BX107633
32	118.4	6.7	378	9	AI222323
33	118.2	6.6	682	12	BM966417
34	116.6	6.6	480	10	AM465382
35	114.2	6.4	556	9	AI342512
36	111.4	6.3	549	10	BF197674
37	110.8	6.2	292	29	CG511844
38	110.6	6.2	540	9	AI582254
39	110.5	6.2	733	14	CA387992
40	108.2	6.1	1185	9	AL552174
41	108.2	6.1	1201	9	AL545270
42	103.8	5.8	587	14	N45367
43	100.2	5.6	654	28	BH074791
44	99.2	5.6	445	10	BF040685
45	97.8	5.5	431	9	AA177901

## ALIGNMENTS

RESULT 1  
LOCUS CF726388 784 bp mRNA linear EST 09-OCT-2003  
DEFINITION UI-M-HBO-ckh-m-15-0-UI-r1 NIH\_BMAP\_HBO Mus musculus cDNA clone  
IMAGE:30547790 5', mRNA sequence.  
CF726388  
CF726388.1 GI:37600556

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Stransberg, Ph.D.  
Email: cga@bcr-remail.nih.gov  
Tissue Procurement: Dr. James Lin University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mouse1.html  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
The following repetitive elements were found in this cDNA

sequence: 280-300, >GC\_richLow\_complexity  
 Seq primer: pyx-5,  
 Location/Qualifiers  
 source  
 1. 784

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30547790"  
 /cissue\_type="whole eye"  
 /dev\_stage="embryo 12.5, 13.5, 14.5 dpc"  
 /lab\_pos="DH10B (T1 phage resistant)"  
 /clone\_lib="NH\_BMAP\_H50"  
 /note="Organ: Eye; Vector: pyx-Asc; Site: 1: Ecor I; Site: 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TATGAGCT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 39.6%; Score 703.6; DB 14; Length 784;  
 Best Local Similarity 94.1%; Pred. No. 4.9e-192;  
 Matches 738; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

122 CTGGAGACAGAGACCTTGGCTGTACAGAGCTGAGTGAAGAAAGACCTTA 181  
 1 CTGGAAACAGAGACCTTGGCTGTACAGAGCTGAGTGAAGAAAGACCTTA 60  
 182 TTCTACTGATGTAGAGACCAAGACCACTGACAGAGCTGAGTGAAGTCTCTC 241  
 61 TTCTACTGATGTAGAGACCAAGACCACTGACAGAGCTGAGTGAAGTCTCTC 120  
 242 TGCAGCTGATCGGTCCATCATTCATGATGCTGCTCATTTCTTAAGTCCAGCGCT 301  
 121 TGCAGCTGATCGGTCCATCATTCATGATGCTGCTCATTTCTTAAGTCCAGCGCT 180  
 302 TGGTGAACCTGGCCGGGAGCTTGGCCGCTTTCTGCGCTTGGGGGCAAAAGACCG 361  
 181 TGGTGAACCTGGCCGGGAGCTTGGCCGCTTTCTGCGCTTGGGGGCAAAAGACCG 240  
 362 ACTTCTGACATTCAGACCTTGAAGAACCCGGGCGGCGCGCGCGCGCGCGCG 421  
 241 ATTTCCTGCGAGTTCAGAACCTTGAAGAACCCGGGCGGCGCGCGCGCGCGCG 300  
 422 ATTACTCTCAAAAACCTATGAGAGATGATGTTGGAAGTGAATGTTGCTGATTAAC 481  
 301 ATTACTCTCAAAAACCTATGAGAGATGATGTTGGAAGTGAATGTTGCTGATTAAC 360  
 482 AGAAAGGCTGCAAGATTGCCAGACCTGATGTTATGCTGTCTCCAAAGGAGAAAG 541  
 361 AGAAAGGCTGCAAGATTGCCAGACCTGATGTTATGCTGTCTCCAAAGGAGAAAG 420  
 542 CAGTGCAGATGATGCTGCTCTTCTTAAGAGAGCAATTCCTCAATCTTACAGATCTCA 601  
 421 CAGTGCAGATGATGCTGCTCTTCTTGAAGAGAGCAATTCCTCAATCTTACAGATCTCA 480  
 602 TATTAACAGCAGAGCTCTTGAACAACCTTATTAATCTTGTGATGCTGTGAGCTCAC 661  
 481 TATTAACAGCAGAGCTCTTGAACAACCTTATTAATCTTGTGATGCTGTGAGCTCAC 540  
 662 TGAATTTGCTTAAAGCACTGCTGATATCCCAATTAATCTTCTGGAAGAGTTCAGTG 721  
 541 TGAATTTGCTTAAAGCACTGCTGATATCCCAATTAATCTTCTGGAAGAGTTCAGTG 600

QY 722 CCTGAGCTGTTGAGTACAGCGCCAGCAAAAAGTACAACTTTCTGAGAACTGGGTA 781  
 DB 601 CCTGAGCTGTTGAGTACAGTACAGTCCAGCAAAAAGTACAACTTTCTGAGAACTGGGTA 660  
 QY 782 ATGAGCCAAATTAATCTATGAGAGCAATGATGAGCGGCGAGTAAATGAGAGCGCTGGAA 841  
 DB 661 ATGAGCCAAATTAATCTATGAGAGCAATGATGAGCGGCGAGTAAATGAGAGCGCTGGAA 720  
 QY 842 AGATTATCATTCAGCTGAGAGAGCTGTTGAGCGCCATTCGAGATTATTCAGAGCCAGCT 901  
 DB 721 AGGA-TTATTCAGCTGAGAGAGCTGTTGAGCGCCATTCGAGATTATTCAGAGCCAGCT 779  
 QY 902 TATA 905  
 DB 780 TGTA 783

RESULT 2  
 LOCUS B0116041  
 DEFINITION  
 ACCSSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
 1 (bases 1 to 628)  
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,  
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,  
 Kono,M., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,  
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,  
 Sano,H., Saeki,D., Shibata,Y., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,P.,  
 Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 Unpublished (2001)  
 On Jun 27, 2000 this sequence version replaced gi:8768609.  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multichannel sequencing. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Kono,M., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,  
 Sugahara,Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Yamazaki,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,  
 Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,  
 Arakawa,T., Ishii,Y. and Hayashizaki,Y.  
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.  
 Funct. Genomics 2 pre, L72-L86 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for  
 further details.

TITLE  
 JOURNAL  
 COMMENT







QY 419 CGGATTAATCTCTCAAAAATGAGATGACATTGTCGAGATGATGTCCTTAGATA 478  
 DB 233 -----GACATTGTCGAGATGATGTCCTTAGACA 263  
 QY 479 AACAGAAAGGCTGCAAGATTGCCAGACACCTGATGATGATGTCCTAGAGAGA 538  
 DB 264 AGCAAAAGGCTGTAGATGTCAGCAGCAGCTGATGTCAGAGCTCCAGAGAGA 323  
 QY 539 AGCAGCTGAGATGATGATGTCCTGATGTCGATGTCGATGTCGATGTCGATGTC 598  
 DB 324 AGCAATCGAGATGATGATGTCCTGATGTCGATGTCGATGTCGATGTCGATGTC 383  
 QY 599 TCATATTAACAGCAGCTGTCAGACAACTTTATTAATCTTCTGATGTCCTGACATCC 658  
 DB 384 TCATATTA-----CAGCTCTAGACAACTTTATTAATCTTCTGATGTCCTGACATCC 439  
 QY 659 ACCTGATATTTGCTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 718  
 DB 440 ACCTGATATTTGCTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 499  
 QY 719 GTGCTGATGATGTCGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 778  
 DB 500 GTGCTGATGATGTCGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 559  
 QY 779 GTATAGGCAAAATTAATCTAGGACATGATGTCGATGTCGATGTCGATGTCGATGTC 838  
 DB 560 GTATAGGCAAAATTAATCTAGGACATGATGTCGATGTCGATGTCGATGTCGATGTC 619  
 QY 839 G 839  
 DB 620 G 620

RESULT 5  
 BM083129 573 bp mRNA linear EST 14-NOV-2001  
 LOCUS fu29f07.y1 Campbell zebrafish ovary Danio rerio cDNA clone  
 DEFINITION IMAGE:5306653 5' similar to TR:Q9QZP8 Q9QZP8 HEPARANASE.; mRNA  
 sequence.  
 accession: BM083129 GI:16930059  
 VERSION EST.  
 KEYWORDS Danio rerio (zebrafish)  
 SOURCE Danio rerio  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 573)  
 Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,  
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theisig, B., Allen, M., Bowers, Y.,  
 Petersen, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,  
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.  
 Washu Zebrafish EST Project 1998  
 Unpublished (1998)  
 CONTACT: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbralist@watson.wustl.edu  
 CDNA Library constructed by library constructed by Invitrogen and  
 donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA).  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 clone distribution: the I.M.A.G.E. Consortium/LML, send email to:  
 info@image.llnl.gov  
 High quality sequence stop: 439.  
 Location/Qualifiers  
 1..573  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"

ORIGIN  
 Query Match 17.28; Score 306; DB 12; Length 573;  
 Best Local Similarity 72.84; Pred. No. 7.5e-77;  
 Matches 409; Conservative 0; Mismatches 150; Indels 3; Gaps 1;  
 /clone="IMAGE:5306653"  
 /sex="female"  
 /dev stage="4-5 months, 1 year and 2 years"  
 /lab host="DH10B (phage-resistant)"  
 /clone\_lib="Campbell zebrafish ovary"  
 /note="Organ: ovary (pooled); Vector: pCMV-Sport6; Site: 1;  
 Note1; Site 2; Salt: whole ovaries collected from zebrafish  
 aged 4-5 months, 1 year and 2 years. Oligo-dT primed,  
 directionally cloned. Average insert size 2 kb. Library  
 constructed by Invitrogen and donated by R. Campbell  
 (Marine Biology Laboratory, Woods Hole, MA)."

RESULT 6  
 AK040471 2173 bp mRNA linear HTC 19-SEP-2003  
 LOCUS Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched  
 DEFINITION library, clone:443010M04 product:heparanase, full insert sequence.  
 accession: AK040471 GI:26333764  
 VERSION AK040471.1  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carmichael, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning

QY 289 AGCTCCAGGCGCTTGATGACCTGGCCCGGAGACTTTCGCGCTTTCGCGCTTCGCGG 348  
 DB 4 ATCTCCAGGCTTTAGTCACTTGACGCGGGACTCTTCGAGCTTATTTAGGTTGG 63  
 QY 349 GGCAGAAAGACCGACTTCTGCACTTCAGACCTGAGAAACCCGCGGAGAAAGCCGCG 408  
 DB 64 GGCAGAAAGACCGACTTCTGCACTTCAGACCTGAGAAACCCGCGGAGAAAGCCGCG 123  
 QY 409 GGGCCGGCCCGGATTTCTATCTCAAAAAGATGAGATGACATTTGTCAGATGATG 468  
 DB 124 ---CTGACCGAGCTATTAAGTCAAAAAGATGAGATGACATTTGTCAGATGATG 180  
 QY 469 GCCTTAGATAACAGAAAGGCTGCAAGATGCGCAGACCCGATGATGATGATGCTGCTC 528  
 DB 181 GCTTGGACAAAGAAAGGCTGCAAGATGCGCAGACCCGATGATGATGATGATGCTGCTC 240  
 QY 529 CAAGGAGAAAGGCTGCAAGATGCGCAGACCCGATGATGATGATGATGATGCTGCTC 588  
 DB 241 CAGGAGGAGAAAGGCTGCAAGATGCGCAGACCCGATGATGATGATGATGATGCTGCTC 300  
 QY 589 TAGAGTATGCTGATTAACAGAGAGCTCTGAGACAACTTATTAATCTTGGATGATG 648  
 DB 301 TAGAGTATGCTGATTAACAGAGAGCTCTGAGACAACTTATTAATCTTGGATGATG 360  
 QY 649 TCTGAGTCAACCTGATATTTGCTGATTAATGCACTGCTGATTAATGCTGCTGCTG 708  
 DB 361 GCAGAGCTGCAACCTGATATTTGCTGATTAATGCACTGCTGATTAATGCTGCTGCTG 420  
 QY 709 AACAGTCTGATGCTGCTGATTAATGCTGATTAATGCACTGCTGATTAATGCTGCTG 768  
 DB 421 AATGCAACAGTCTGCTGATTAATGCTGATTAATGCACTGCTGATTAATGCTGCTGCTG 480  
 QY 769 TGGAGACTGGTATTAAGCCCAATTAATGCACTGATTAATGCACTGATTAATGCACTG 828  
 DB 481 TGGAGACTGGTATTAAGCCCAATTAATGCACTGATTAATGCACTGATTAATGCACTG 540  
 QY 829 AGCGAGTGGAGAAAGATTACA 850  
 DB 541 TCTGCTGCTGCTGAGATTACA 562

JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komono, H., Akiyama, Y., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
MEDLINE 11076861  
PUBMED 11076861  
REFERENCE 5  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
MEDLINE 11076861  
PUBMED 11076861  
REFERENCE 6  
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanekawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, N., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, T., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenryo-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
FEATURES  
source Location/Qualifiers  
1..2173  
/organism="Mus musculus"  
/mol\_type="cDNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:A430101M04"  
/db\_xref="MGI:2405157"  
/db\_xref="Ensembl:10090"  
/clone="A430101M04"

ORIGIN  
Query Match 16.0%; Score 284.2; DB 11; Length 2173;  
Best Local Similarity 52.9%; Pred. 4e-70;  
Matches 678; Conservative 0; Mismatches 573; Indels 30; Gaps 2;  
CDS  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
91..1698  
/note="unnamed protein product; hepatanase  
(MGI:1343124, GB|AA047943, evidence: BLASTN, 99%,  
match=480)  
putative"  
/codon\_start=1  
/protein\_id="BAC30600.1"  
/db\_xref="MGI:2633765"  
/translation="MRLILIMVGLALAGAPAGAPAPDDVVDLEPTKRLRSV  
SPESLITIDASLADPRLFTLFGSPRLALRGLSPALRFGKRTDLIDPDEP  
TSEERLYMSQVHDIQSEPSAAVRLQVWPQILDLREDOYKFKSTYRS  
SVMLVSPAKSGSLDIPELNLRLTPDRNMSNAQLDDYCSKGNISWELNEP  
NSFWKXALHILIDQLQEDFVEHLKQSAFONAKLYDPDIPGPKVLLRFLK  
AGEVEIDSLWHYVYNGRIATKEDPLSDVDVDFILSYOKIKYKEITPEKRWLG  
ETSAVAGGAPLSTNFAAGFMWLDLGLSAQMGIEVMROYFPGANVHVDENEP  
LDPYVLSILFKLVGRVILSRKGPDSKLELVYLCNTVYHRYROEGDITLVYVNH  
NVRKHVPPPLFRKPVDTYLLKPSGPDGLSKSVQNLQIIMVDEDTLPALTEKPL  
PAGSALSLPFSYGFVIRNAKIACI"  
474 AGATTAACAGAAAGGCTGCMAAGTTGCCAGACCTGATGTTATGCTGCTCCAAAG 533  
429 AGTCACACATGATATTTGGCAGCTCGACGGCTCTGCTGGCGTGTGAGGAATCTCCA 488  
534 GGAGAAGCAGCTCAGATGATGCTGCTCTCAAGAGCAATCTCAATCTACAG 593  
489 GGTGGAATGGCCCTTCAGAGAGCTGTTGCTCCAGAGCACTACCAAGAGTTCA 548  
594 TAACTCATTTAACAAGCCAGTCTCTAGCAAACTTTAACTTTGCTGATTCTCG 653  
549 GAACAGACCTACTCAAGAAAGCTCAGTGAACATGCTCTCAAGTTTGCAGAGCTCG 608  
654 ACTCCACCTGATTTTGTCTTAATGACATGCTGCTTAATCCAAATCTCCGGAACG 713  
609 GTTAGACTATCTTGTGTAAATGCTTACATCAACACCCAGATTACGTTGAACAG 668  
714 TTCTAGTCCCTGAGTCTGTGAGTACAGCCGACCAAAAGTACACATTTCTTGG 773  
669 CTCCAAGCCAGCTTCTCTTACATGCTCTTCCAAAGGTTAATACATCTCTGG 728  
774 ACTGGTAAAGACCAATTAATCTGACATGACATGACGCGGACGTAATGGCAGCA 833  
729 ACTGGCAATGAGCCCAACGTTTCTGGAAGAAAGCTCACTATTCATGATGGTTGCA 788  
834 GTTGGGAAGATTATCATCCAGCTGAAGAGCTGTTGCAAGCCATCCGATTATTCAG 893  
789 GTTAGGAAGAGCTTGTGAGTGTGATTAATCTTACA--AAGTCAAGCTTCCAAA 845  
894 AGCCAGCTTAATGAGCTTAATTTGGGGGCGGAGAGAAAGTATGATGCTCCTTAA 953  
846 TGCAAAACCTAATGCTGCTGACATGCTGACGCTCAAGGAGAGACATTAATGCTGAG 905  
954 TGATTCATAGAGTGGCAGAGAGTACAGTATGACATGATTCAGCAATTCATCAAT 1013  
906 GAGTTTCTGAAGGCTGGGAGGAAGATGATGATCTCTTAATGATGATCACTATTAAT 965  
1014 TGATGGCGGGGTGCTCAAGTATGACATTCCTGAAAATCTGCTTTAGACACTGTC 1073  
966 GAATGACGATGCTACCAAGAAATTTTGAAGCTGATGATGCTGACACTTTAT 1025  
1074 TGACCAATAGAGAAATTCAGAAATGTTAATATACATCTCCAGAAAGAGATTGG 1133  
1026 TCTCTGTGCAAAAATTTGAAAGTCACTAAAGATCAACCTGGCAAGAGTCTGG 1085  
1134 GCTTGAAGTGTGTGACCACTCAGCTGAGGCAAAACATCTATCCGATTCCTATGC 1193

Db 1086 GTTGGAGAGACGAGCTACGTTACGCTGGCGGTGACACCTTGTCTCCACACCTTTGC 1145  
 QY 1194 TGCAGATTCTTATGTTGTAACAATTAGAAATGCTGGCCAAATCAGGGCTATGATGTGT 1253  
 Db 1146 AGCTGGCTTATGCTGCTGATTAATTTGGGCTGTCCAGCCAGATGGGCTAGAAATGCT 1205  
 QY 1254 GATACGGCACTATTGTTTGGACCATATACATCACTCGTGGACAGAAATTTAAACC 1313  
 Db 1206 GATGAGCGAGGTTCTTGGAGACGCACTACCTTAGTGAATTAATTTGAGCC 1265  
 QY 1314 ATTACGAGTACTGCT 1373  
 Db 1266 TTACTGATGATGCTGCT 1325  
 QY 1374 TGTGATGATGCTGCTGCT 1433  
 Db 1326 GTCAAGATGAAAGGCCAGA-----CAGAGCAACT 1358  
 QY 1434 AAGATTATGCT 1493  
 Db 1359 CCGAGTATCT 1418  
 QY 1494 ACTTTTATCT 1553  
 Db 1419 TGTGATGCT 1478  
 QY 1554 AGAAGAGCTGTTTACCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1613  
 Db 1479 CAGAGAACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1538  
 QY 1614 GTGATGCT 1673  
 Db 1539 ATCTGCTCACTGAGCGCTCAATTTCTGAAGTGTGATGATGATGATGATGATGAT 1598  
 QY 1674 GAGCCCGCCCT 1733  
 Db 1599 GAGAGAAACCT 1658  
 QY 1734 TTATGCTGCTCAAGATGCTCA 1754  
 Db 1659 TTTTGCATAGAAATGCCAA 1679

RESULT 7 640 bp DNA linear GSS 27-SEP-2003  
 CE365946  
 LOCUS  
 DEFINITION  
 tigr-gss-dog-17000362155477 Dog Library Canis familiaris genomic.  
 genomic survey sequence.  
 CE365946  
 VERSION  
 CE365946.1 GI:36583778  
 KEYWORDS  
 GSS.  
 SOURCE  
 ORGANISM  
 Canis familiaris (dog)  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 1 (bases 1 to 640)  
 Kirchner, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,  
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
 Venter, J.C.  
 The dog genome: survey sequencing and comparative analysis  
 Science 301 (5641), 1898-1903 (2003)  
 14512627  
 Contact: Kirchner EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirchner@tigr.org  
 Class: shotgun.  
 Location/Qualifiers  
 1..640

FEATURES  
 source  
 1..640

Query Match 15.4%; Score 274.4; DB 29; Length 640;  
 Best Local Similarity 96.2%; Pred. No. 1.2e-67;  
 Matches 281; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="standard poolie"  
 /db\_xref="taxon:9615"  
 /clone\_1db="Dog Library"  
 /note="Site 1: BstXI; Libraries were prepared from  
 peripheral blood"

ORIGIN  
 1 ATGAGGATGCTTTGCTTCT 60  
 174 ATGAGGATGCTTTGCTTCT 233  
 61 TGCCTAGCCCGGGGGGCT 120  
 234 TGCCTAGCCCGGGGGGCT 293  
 121 GCTGAGACAGAGACCTTGTCTGTAGACAGAGCTGACAGTTGAGAGAAAGACCTGT 180  
 294 GTTGAAGACAGAGACCTTGTCTGTAGACAGAGCTGACAGTTGAGAGAAAGACCTGT 353  
 181 ATTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 354 ATTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413  
 241 CTGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 292  
 414 CTGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 465

RESULT 8 351 bp mRNA linear EST 16-JUN-1998  
 A1019269  
 LOCUS  
 DEFINITION  
 ub22a07.r1 Soares\_thymus\_2NbwT Mus musculus cDNA clone  
 IMAGE:1378452 5', mRNA Sequence.  
 A1019269  
 VERSION  
 A1019269.1 GI:3233605  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 351)  
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisler, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 The WashU-HMNI Mouse EST Project  
 Unpublished (1996)  
 Contact: Marra M/Mouse EST Project  
 WashU-HMNI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LBNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:900920  
 Seq primer: -28n3 rev2 ET from Amersham  
 High quality sequence stop: 337.  
 Location/Qualifiers  
 1..351  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone\_1db="IMAGE:1378452"

FEATURES  
 source  
 1..351



QY 474 AGATTAACAGAAAGGCTGCAAGATTGCCAGACACCTTGATGTTATGCTGTGCTCCAAAG 533  
 DB 455 AGTCAACCATGATATTTGACAGCTGAGCCGGTCTCTGCGGGGTGAGAAATCTCA 514  
 QY 534 GAGAGAGGACAGTCAATGATCTGTCTTCTTAAGAGACATCTCCATCTACTTACAG 593  
 DB 515 GGTGGAATGGCCCTTCCAGAGAGCTGTCTGCTCCAGAGACATCAAAAAGAGTTCAA 574  
 QY 594 TAATCTCATATTAAAGCAGAGCTCTAGACAACTTTAATCTTGTGATGCTGTG 653  
 DB 575 GAAAGACACCTACTAGAAAGCTCAGTGGACATGCTCTACAGTTTGGCAAGTGTGGG 634  
 QY 654 ACTCACTGATATTGCTCTTAATGACACTGCTGTGTAATCCCAATACCTCTGAGAC 713  
 DB 635 GTTAACCTGATCTTGTGCTTAATGCTTAATGCAACCCAGCTTACGATGAGACAG 694  
 QY 714 TTCTAGTCCCTGAGCTGTGTAAGACAGCCGCAAAAAGTACACATTTCTTGGGA 773  
 DB 695 CTCACAGCCGCAAGCTTCTCTGACTACTGCTCTTCCAAAGGTTTAAACATCTCTGGA 754  
 QY 774 ACTGGATTAAGACCAATTAATCTATGGAACCATGATGCGCGGAGTAAATGCGACCA 833  
 DB 755 ACTGGGCAATGAGCCCAACAGTTCTGGAAGAAAGCTCAATTCATGATGAGTTGCA 814  
 QY 834 GTTGGGAAAGATTCATCAGCTGAAGAGCTGTGACAGCCATCCGATTTATTCAG 893  
 DB 815 GTTGAAGAAAGCTTTGTGAGATTGATTAATTTCTACA--AGTCAAGCTTTCAAA 871  
 QY 894 AGCCAGCTTAATGAGCCCTAATATTGAGGCGCCGAGAGAAATGTCAGCCCTCTAGA 953  
 DB 872 TCAGAAACCTAATGAGCTGATCGATCGATCAGCTCGAGGAAAGACAGTTAACTGTCAG 931  
 QY 954 TGGATTCAGAGAGTGGGAGAGAGTACAGTATGACATGTTCTGCGCAACCTGTATAT 1013  
 DB 932 GAGTTCTCGAAGGCTGCGGAGAAAGTATGATCTCTTACATGATGCACTATTAAT 991  
 QY 1014 TGAATGCGCGGAGTGTCAAGTATGAGACTTCTGAAAACCTGCTGTAGACACACTCTC 1073  
 DB 992 GATGAGAGCATGCTGACCAAGAAAGATTTTCTGAGCTGTGATGCTGAGACACTTTAT 1051  
 QY 1074 TGAACGATTAAGAAATTCAGAAAGTGTTAATCATACACTCAGAGAAAGAAATTTG 1133  
 DB 1052 TCTCTCTGCAAAAATTTCTAAGGTCACTAAGAGATCAACACTGCGAAGAAAGTGTG 1111  
 QY 1134 GCTTGAAGTGTGTGACCACTCAGCTGAGGAGCAAAAACATCTATCCGATTCTATGC 1193  
 DB 1112 GTTGGAGAGAGAGAGTCACTTACGATGCGGTGCAACCTGCTGCCAACACTTTGC 1171  
 QY 1194 TCGAGAGATTTATGATGTAACAATTAGAAATGCTGCGCAATCAGGCAATTGATGCT 1253  
 DB 1172 AACTGCTTATGTGCTGATGTAATGAGGCTGTCAAGCCAGATGAGGAGATGAGTGT 1231  
 QY 1254 GATAGGAGCACTATTGTTTGAACCATGATATCAATCACTGCTGAGACCAAAATTTTAA 1313  
 DB 1232 GATGAGGAGAGTGTCTTCTGAGAGAGCACTACATTAAGTGAATGAAAATTTGAGCC 1291  
 QY 1314 ATTACCAAGCTACTGCTCTCTCTCTTCAAGGCGCTGATGCGCCCAAAATCTTGGC 1373  
 DB 1292 TTACTCTGATTAAGTGTCTCTCTCTTCAAGAACTGTGAGGCTCCAGAGGCTTTACT 1351  
 QY 1374 TGTGATGAGTGTGGGCTCAAGGAGGACGAGGCTGCGAGTATCCGGGACAAACT 1433  
 DB 1352 GTCAAGAGTGAAGGCCCCA-----CAGAGAGAACT 1384  
 QY 1434 AAGGATTTATGCTCACTGACAAACCAACCAACCAATACGTTCTGTGGTTCATTAAC 1493  
 DB 1385 CCGAGGTATCTCACTGACATAAGTATACACCACGATATACGAGAGAGATCTAAC 1444  
 QY 1494 ACTTTTATCATCACTGATGATCAAGAAAGAAATCAAGCTGGGCTGGAGACTCTAG 1553  
 DB 1445 TGTGATGCTGAACTCTCAATATGTCAACAAAGACTTGAAGATACCGCTTCGTTGT 1504

QY 1554 AGCAAGCTGCTTACCAGTACTGCTGACAGCCCTATGCGGACAGAGGCTTAAGTCCA 1613  
 DB 1505 CAGAAACCAAGTGAATACCTTCTGAAGCCTTCGGGCGGATGATTTCTTCCA 1564  
 QY 1614 GTCAAT 1619  
 DB 1565 GTCAAT 1570

RESULT 10  
 B1661209 321 bp mRNA linear EST 12-SEP-2001  
 LOCUS 60304542F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:5350211 5'  
 DEFINITION mRNA sequence.  
 B1661209  
 VERSION B1661209.1 GI:15575445  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 321)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch  
 Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.lnl.gov>  
 Plate: LHAM1880 row: C column: 12  
 High quality sequence stop: 321.  
 Location/Qualifiers  
 1..321  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NMRI"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5350211"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_1ib="NCI CGAP Mam4"  
 /note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; Salt;  
 Site: 2; Not: Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Priscilla Furch,  
 NIH Reference for transgenic model: Li et al., Cell Growth  
 and Differentiation 7, 3-11 (1996)."  
 ORIGIN  
 Query Match 13.6%; Score 241.2; DB 12; Length 321;  
 Best Local Similarity 90.2%; Pred. No. 3.1e-58; Indels 0; Gaps 0;  
 Matches 258; Conservative 0; Mismatches 28;

QY 1406 GGCCTGCGCAGTATCCGGAGCAAACTAAGATTATGCTCACTGACAAACCAACACA 1465  
 DB 36 GGCCAGGAAGATCATCCAGACAACTAAGATCTATGCTCACTGACAAACCAACACA 95  
 QY 1466 ACCCAACTACGTTGCGGTCCATTACCTTTTATCATCACTTGCATCGATCAAGAA 1525  
 DB 96 ACCCAACTATGTCGGGCTCTATACACTTTTATCACTCAACCTACATGATCAAGAA 155  
 QY 1526 AGAAATCAAGCTGCGGAGCTCTCAGAGACAAGCTGTTCAACCACTGCTGCTCAG 1585  
 DB 156 AGAAATCAAGCTGCGGAGCTCTCAGAGACAAGCTGTTGACCAAGTATCTGCTCAG 215  
 QY 1586 CTAATGCGAGAGGCTTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1645

Db	216	CCATGTGGGACGAGGAAGCCTTAAGTCCAAGTCAGTGCAACTCATCATGAGCGAGCCTTAGTGA	275
QY	1646	TGTGTGTGCAGCAGAGGAGACCTCCAGATTGAGAGCCCCGCCCTTGC	1691
Db	276	TGTGTGTGCAGCAGGAGACCTTCAGAACTGAAAGCCCCGCCGTGCG	321
RESULT 11			
LOCUS	BU474796		
DEFINITION	BU474796	838 bp	mRNA linear EST 30-NOV-2002
	603765360orf1 CSEORBN21 Gallus gallus CDNA clone ChST865g16 5', mRNA sequence.		
ACCESSION	BU474796		
VERSION	BU474796.1	GI:25968373	
KEYWORDS	EST,		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Atherinomorpha; Aves; Neognathae; Galliformes; Phasianidae;		
	Phasianinae; Gallus.		
REFERENCE	1 (bases 1 to 838).		
AUTHORS	Boatman,F.E., Sanz-Ezquerro,J., Overton,I.W., Burt,D.W., Bosch,E., Fong,W.T., Tickler,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.		
TITLE	A Comprehensive Collection of Chicken cDNAs		
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)		
MEDLINE	22335534		
PUBMED	12445392		
COMMENT	Contact: Simon Hubbard		

**FEATURES**  
 Source  
 1. 838  
 Location/Qualifiers  
 Email: Simon.Hubbard@umist.ac.uk.  
 Tel: 01612008930  
 Fax: 01612350409  
 PO Box 88, Manchester, M60 1QD, UK  
 (UMIST)  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology

```

/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="CHEST655g16"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_idb="CSE0REN21"
/note="Organ: Ovary; Vector: pluescript II KS(+); Site_1
ECORI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntend, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994): 91: 9228-9232 and Bionaldo et al., Genome Research
(1996): 791, except that a significantly longer
reannealing hybridization was used."

```

Query Match	12.0%	Score 212.6	DB 13	Length 838
Best Local Similarity	79.7%	Pred. No. 1.2e-49		
Matches 251	Conservative 0	Mismatches 64	Indels 0	Gaps 0
QY	1005	TTGCTACTTGTATGCGCGGGTGTCAAGTATATGACTTCTCGAAAACCTCGCTGTTAGA	106	
Db	93	TTACTACTCATGATGGCCAGATGGCCAAAGTGACTACTTCCTGAAAACACGGCTCTCGA	152	
QY	1065	CACACTCTCTGACCGATTAGAAAATTGAGAAATGGTTAATACATACATCCAGGAAA	112	

Db	153	CACACTGTCACACGAGATCAGGAAAAATTCAGAAAAATTGTGAACCCATACACACCGGGCGAA	212
Qy	1125	GAAATTTGGGTTGAAGGTGTGTGACCAACTGACGTGAGAGGACAAACAAATTATCCGA	1184
Db	213	GAAATCTGGCTGAGAAAGCGTTTGTGCTACTCTGTCTGAGGACATGAACAACTCTCCGA	272
Qy	1185	TTCTCATGCTGACGAAATTTCTTATGTGTGAACACTTTTAGAAATGCTGGCCATCAAGGCAT	1244
Db	273	TTCTTATGCGGCGCGGGTTCGTGTGTGCGAACACCGCTGGCTTGTGTGGCCACCCAAAGGAT	332
Qy	1245	TGATGTGCTGTAAGGAGCATATTTTGTGACATGATATCAATACCTCGTGGACCGAGA	1304
Db	333	TGAATGATGTGTGCGGCACTTCTTCTTGACCAATGGCCACACCACTGTGTGACACAGAA	392
Qy	1305	TTTTAACCCATTACC	1319
Db	393	CTTCAACCCCTGGC	407

RESULT	12
CG611539	
LOCUS	
DEFINITION	CG611539 483 bp DNA linear GSS 02-OCT-2003
OS	OS1296636 Mus musculus 12SV/Ev Mus musculus genomic clone
ACCSSION	CG611539
VERSION	CG611539.1 GI:37435388
KEYWORDS	GSS.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 483)	Zambrowicz, B. P., Abuhin, A., Ramirez-Solis, R., Richter, L. J., Piggott, J., Beltrandelio, H., Butron, E. C., Huang, W., Finch, R. A., Ridgell, C. J., Gupta, A., Hansen, G., Hu, Y., Edwards, J., Finch, C., Key, B. W., Urz, Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markschl, D., Payne, R., Poter, D. G., Qian, N., Shah, Y., Schrick, J., Shi, Z.-Z., Sparks, M. J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C., and Sands, A. T.	Mhl kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention	Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)	Contact: Zambrowicz BP OmniBank

```

FEATURES
  source
    Lexicon Genetics Incorporated
    4000 Research Forest Drive, The Woodlands, TX 77381, USA
    Email: materials@lexgen.com
    Gene trap sequence tag generated by 3' RACE from mouse ES cells as
    described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)
    Class: Gene Trap.
    Location/Qualifiers
      1..483
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="129SV/Ev"
        /dd_xref="taxon:10090"
        /clone="OST936636"
        /cell_type="embryonic stem cell"
        /clone_1tb="Mus musculus 129SV/Ev"

```

Query Match	11.4%	Score 202.2;	DB 29;	Length 483;
Best Local Similarity	87.0%;	Fred. No. 8.3e-47;		
Matches 241;	Conservative	0;	Mismatches 34;	Indels 2;
				Gaps 2
QY	1340	TCTACAAGCGGCTGATCGGCCCAAGCTTGCGTGCATGTGTGCTGGGCTCCAGCGGA	1399	
Db	1	TTTAAAGAGCTGTGATTGGCCCCAAGTCTCGGCGAGNCAGATGACTGGGCTCCAGCGGA	60	
QY	1400	AGCCACGGGCTGGCGGAGTGTATCGGCGACAACTAAGATTATGCTCACTGCACAAAC	1455	
Db	61	AGCCTCAGCCAGGAAGAAGTCAATCCAGACAAACTTNGAATCTATGCTCACTGCACAAAC	120	

QY 1460 ACCCAACCAACACAGT-TCGGGGGTCATTACA-CTTTTATACCACTTGCAATCG 1517  
Db 121 ACCCAACCAACACATATATCCCGGGATCTATACACCTTTTCAATCAACCACTACATCG 180  
QY 1518 ATCAAGAAAGAAAATCAAGCTGCGTGGACTCTCAGACACAAAGCTGTCCACCAATCT 1577  
Db 181 ATCAAGAAAGAAAATCAACACTGGCTGGGACTCTGAGGACAAAGCTTGTGACACAGTACT 240  
QY 1578 GCTGACGCTTATGGGACGAGGAGGCTTAAAGTCCAG 1614  
Db 241 GCTGACGCTTATGGGACGAGGAGGCTTAAAGTCCAG 277

RESULT 13	AL552151/c	1156 bp	mRNA	linear	EST 31-MAY-2003
LOCUS	AL552151.1				
DEFINITION	AL552151 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSOD1059YNT5 3-PRIME, mRNA sequence.				
ACCESSION	AL552151				
VERSION	AL552151.2	GI:31273967			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1156)				
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12890775.				

Genoscope - Centre National de Séquençage  
BP 191 91006 EVRY cedex - France  
Email: [seger@genoscope.cns.fr](mailto:seger@genoscope.cns.fr)  
Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 2469.r For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSD01059CG08NP1&cluster=2469.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSD01059CG08NP1&cluster=2469.r). Contact :  
Feng Liang Chuaier: [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
Paradey Avenue Genoscope sequence ID : CSD01059CG08NP1.

FEATURES	Location/Qualifiers
source	1. .1156

```

1..1156
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1059Y15"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	10.8%;	Score 191.6;	DB 9;	Length 1156;
Best Local Similarity	51.6%;	Pred. No. 1.8e-43;		
Matches 525; Conservative	3;	Mismatches 458;	Indels 31;	Gaps 3

OY	750	CAAAAAGTACAACTTTCTTTGGAGACTGGGTATATGAGCCAAATATCTTGGACCATGCA	809
	986	CCAAGGGTATATTAATTTCTTGGACACTGAGCAATGACCTAA-CAGTTCCTTTAAGAAAGC	928
Db	810	TGGCCGGGCAGTAAATGGAGCCAGTGGGAAAGATTACATCCAGCTGAAGAGCCGTT	869
OY	927	TGATATTTTATCAATGGGTGCGAGTTAGAGAGATTTTATTTCAATTGCTAATAC--T	871
Db	870	GCAACCACATCCGATTTTATCCAGAGCCAGCTTATATGGCCCTTAATTGGGCGGCGAG	928
OY	870	TCTAAGAAATGCCACCTTAAAAATGCAAAACCTATGGTCTGATGGTCAAGCCTCG	811
Db	930	GAGAGATGATGCGCCCTCTAGATGATTCATGAAGTGGCAGGAGATGACATGATATC	985
OY			

Db 810 AAGAAAGACGGCTAAGATGCTGAAGAGCTTCCTGAAAGCGTGGTGAAGAAATGATGATTCC 751

Qy 990 AATTACCTGGGGAACATTTGCTACATTTGATGGCCGGGAGCTCAAGATGATGGACTTCCTGAA 1049

Db 750 AATTACATGGGACATCTATCTATTGATGAGCGGACTGCTACAGGGAAGATTTCTTAA 691

Qy 1050 AACTGCGCTGTTAAGACACACTCTGTGACAGATTAGAAAATTCAGAAAGGTGTTATAC 1109

Db 690 CCTGATGATTTGGCAATTTTATTTATTCATCTGTGCAAAAAGTTTCCAGTGGTTGAGAG 631

Qy 1110 ATACACTCCAGAAAGAAAGATTTGGCTGTAAGGTGTGGACCACTCAGCTGGAGGAC 1168

Db 630 CACGAGGCTGGCAGAAAGGTCTGGTTAGGAGAAACAGCTCTGCATATGAGGGGGAGC 571

Qy 1170 AAACATCATTCGATTCCTATGCTGACAGATTCTTATGGTTGAACATTTAGATGCT 1229

Db 570 GCCCTTGCTATCCACACTTTGAGCTGGCTTTATGTGGCTGAATAATTGGCGWCTC 511

Qy 1230 GGCCAAATCAGGCAATTGATGTGTGATACGGCACTATTTTGTACATGATACATCA 1289

Db 510 AGCCCGAATGGGAATAGAAAGTGGATGAGAGCAAGATTCTTTGAGCAGGAAACTTACA 451

Qy 1290 CCTCGTGACAGAAATTTTAAACCATTACAGACCTACTGGCTCTCTCTCTTACAACG 1349

Db 450 TTTATGTGATGAATCTTCGATCTTTACTGATTAATGGCTATCTCTTCTGTCAAGA 391

Qy 1350 CCTGATCGGCCCAAAAGCTTGGCTGTGATGGCTGGCTGGCTCCAGCGGAGCCAGGC 1408

Db 390 ATTGGTGGGCACAAAGTGTTAATGGCAACGTGCAAGGTTCAAAGAGAGG----- 339

Qy 1410 TGGCCGATGATCCGGGACAAACTAAGATTTATGTCTCACTGACAAACACACACAGCA 1468

Db 338 -----AAGCTTGAGTAACTTTCATTGGACAAACACTGACAAATCC 298

Qy 1470 CAACTACGTTCTGGGTCATATTACATCTTTTATCATCACTTGCAATCATCAAGAAAGAA 1528

Db 297 AAGGATTAAGAAGAGATTTAACTCTGTATGCAATAAACCTCATATATGTCAACACGA 238

Qy 1530 AATCAAGCTGGCTGGGAATCTCAAGACAGCTGTTCACCACTACTCTGTGCAACCTCA 1588

Db 237 CTTCGGGTTACCTTATCTTTTCTTAAACAAGCAATGGAATTAATACCTTCTAAAGACTTT 178

Qy 1590 TGGGACAGAGGGCTTAAAGTCCAAAGTCAAGTCACTGAATGGCCAGCCCTTATGATAGT 1648

Db 177 GGGAGCTCATGATTAATCTTCCAAATCTGTCAACTCAATAGGTCTAACTCTAAAGTGT 118

Qy 1650 GGAAGAGGGGACCTTCCAGAAATTGAAGCCCGCCCTTCTGGGCGCGCGGACATTTGT 1708

Db 117 GGAATGATCAAACTTTGCACCTTTAAATGAAAAACCTTCCGGCAGGAAGTTCACTGGG 58

Qy 1710 CATCCCTCAGTCAACATGGGCTTTATATGAGTCAAAATGTCAATGCTTTGGCGT 1766

Db 57 CTTCGACAGTATCTATATAGTTTTTTTGTGATTAAGAAATGCCAAAGTTGCTGCTKG 1

[illegible]

Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.  
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention  
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
 Contact: Zambrowicz BP

OmniBank  
 Lexicon Genetics Incorporated  
 4000 Research Forest Drive, The Woodlands, TX 77381, USA  
 Email: materials@lexgen.com  
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)  
 Class: Gene Trap.  
 Location/Qualifiers  
 1. 334  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129SV/ev"  
 /db\_xref="taxon:10090"  
 /clone="OST419746"  
 /cell\_type="embryonic stem cell"  
 /clone\_id="Mus musculus 129SV/ev"

## ORIGIN

Query Match 10.5%; Score 187.6; DB 29; Length 334;  
 Best Local Similarity 82.6%; Pred. No. 1,1e-42;  
 Matches 242; Conservative 0; Mismatches 47; Indels 4; Gaps 3;

QY 1325 ACTGGCTCTCTCTCTCTTACAGCGCTGATCGCCCAAGCTTGGCTGTCATGTGG 1384  
 Db 1 ATGGCTCTCTCTCTCTTCAAGCGTGTGATGGCCCAAGCTTGGCGATGATG 59  
 QY 1385 C-TGGGCTCAGCGAAGCCAGCGCTGCGCGAGTATCGGGAACAATTAAGATTAT 1443  
 Db 60 CGTGGCTCAGCGAAGCGCTCGCCGGAAGTATCGCGAACAATTAAGATTAT 119  
 QY 1444 GCTCAGTCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1501  
 Db 120 GCTCAGTCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 179  
 QY 1502 TCATCAACTGCATCATCAAGAAAGAAATCAAGCTGAGGCTCTCAGAGCAAGC 1561  
 Db 180 TCATCAACTGCATCATCAAGAAAGAAATCAAGCTGAGGCTCTCAGAGCAAGC 239  
 QY 1562 TGGTTCACCAATCACTGCTGAGCGCTTATGGGAGAGGAGGCTTAAGTCAAG 1614  
 Db 240 TTGTGCGCCATATGCTGAGCGCTTATGGGAGAGGAGGCTTAAGTCAAG 292

## RESULT 15

AL545232/c 1200 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL545232 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION AL545232  
 ACCESSION AL545232  
 VERSION AL545232.2 GI:31267068  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1200)  
 I, W.B., Gruber, C., Jesse, J. and Polyes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12877713.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2469.r For

more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1028DC02NP1&cluster=2469.r>. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
 Faraday Avenue Genoscope sequence ID : CS0D1028DC02NP1.  
 Location/Qualifiers  
 1. 1200  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1028YF04"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 9.6%; Score 170.8; DB 9; Length 1200;  
 Best Local Similarity 51.6%; Pred. No. 2e-37;  
 Matches 431; Conservative 3; Mismatches 375; Indels 27; Gaps 1;

QY 902 TATATGGCCCTATATTTGGGCGCGGAGAGATGTCATGCGCCCTCTAGATGATTC 961  
 Db 809 TCTATGGCTGATGTTGGTCAAGCGGAAAGAGCGGAAATGCTGAGAGCTTC 750  
 QY 962 TGAAGTGGCAGAGAGTACATGATGATGATGATGATGATGATGATGATGATGATG 1021  
 Db 749 TGAAGTGGTGGAGAGTATGATGATGATGATGATGATGATGATGATGATGATG 650  
 QY 1022 GGGTGTCAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1081  
 Db 689 GGAATGCTACAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATG 630  
 QY 1082 TTAGGAAATTCAGAAAGTGGTATATATATATATATATATATATATATATATAT 1141  
 Db 629 TGCAAAAGTTTCAGAGTGTGTAGAGACACAGCGCTGGCAAGAGTGTGTAGAG 570  
 QY 1142 GTGTGTGACACCTCAGCTGAGGAGGACAAACATATATATATATATATATATAT 1201  
 Db 569 AAAGAGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 510  
 QY 1202 TCTATGCTGATCACTTATGAGATGCTGGCCATCAGGCAATGATGATGATGATG 1261  
 Db 509 TTATGTGCTGATCACTTATGAGATGCTGGCCATCAGGCAATGATGATGATGATG 450  
 QY 1262 ACTCATTTTTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1321  
 Db 449 AAGATATCTTGGAGCAGAGAACTACATTTAGTGAATGAAGAACTTGAATCTT 390  
 QY 1322 ACTATGCT 1381  
 Db 389 ATTATGCT 330  
 QY 1382 TGGCTGGGCTCCAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1441  
 Db 329 TGCAAGTTCAGAGAGAGG-----AAGCTTCAGAT 297  
 QY 1442 ATGCTCATGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1501  
 Db 296 ACCCTCATGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 237  
 QY 1502 TCATCAACTGCATCATCAAGAAAGAAATCAAGCTGAGGCTCTCAGAGCAAGC 1561  
 Db 236 CCAATTAACCTCATATATGATCAAGATCTGGGTTACCTTATCTTTTCTTACAGC 177  
 QY 1562 TGGTTCACCAATCACTGCTGAGCGCTTATGGGAGAGGAGGAGGAGGAGGAGGAG 1621  
 Db 176 AAGTGAATTAATATCTTCTAAGACCTTGGAGACCTCATGATTAATCTTCAAAATG 117  
 QY 1622 AATGAATGGCAGAGCGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1681



Fri May 14 13:01:54 2004

us-10-088-676-1.rst

Page 13

Db 116 AACCTCATGAGTCTAACTTAAGATGATGATCAAACTTGCCACCTTTAATGAAA 57  
QY 1682 GCCCCTTCGGGCGGCGGACATGTCATCCCTCCAGTCAACATGAGCTTTAT 1737  
Db 56 AACCTCTCGGCGCAGGAAGTTCACCTGGGCTTGCCAGCTTTCTCATATATTTT 1

Search completed: May 13, 2004, 15:37:47  
Job time : 4747 secs





QY 421 HSPFHGYNHLYDONENPLPDYWLSLYKRLIGPKYLAHVAGLORRPPRGRVIRDKLRI 480  
 DB 421 HSPFHGYNHLYDONENPLPDYWLSLYKRLIGPKYLAHVAGLORRPPRGRVIRDKLRI 480  
 QY 481 YACCTNHNHNHYVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGELKSKSV 540  
 DB 481 YACCTNHNHNHYVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGELKSKSV 540  
 QY 541 QUNGQPLVWVDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592  
 DB 541 QUNGQPLVWVDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592  
 RESULT 3  
 ID AAU07424 standard; protein; 592 AA.  
 XX AAU07424;  
 AC  
 XX 18-DEC-2001 (first entry)  
 DT  
 XX  
 DE Human heparanase-like protein splice variant #1.  
 XX  
 KM Human: immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
 KM antiproliferative; cardiact; vasotropic; cerebroprotective; nootropic;  
 KM neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 KM extracellular matrix; ECM; autoimmune disease; rheumatoid arthritis;  
 KM hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KM cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection;  
 KM nervous system disorder; Alzheimer's disease; ocular disorder; sunburn;  
 KM wound healing; food additive; heparanase.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179253-A1.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 11-APR-2001; 2001MO-US011643.  
 XX  
 XX 18-APR-2000; 2000US-0198123P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Fiscella M, Shi Y, Edner R, Ruben SM;  
 XX  
 DR WPI; 2001-611720/70.  
 XX  
 DR N-PSDS; AAS13848.  
 XX  
 PT New nucleic acids encoding extracellular matrix polypeptides, for  
 PT diagnosing, treating, preventing or ameliorating human disorders and  
 PT disease, such as, autoimmune, hyperproliferative or cardiovascular  
 PT disorders.  
 XX  
 PS Disclosure; Page 14; 308bp; English.  
 XX  
 CC The invention relates to novel isolated polynucleotides (I) encoding  
 CC extracellular matrix (ECM) polypeptides. (I) and a polypeptide encoded by  
 CC (I) are used to prevent, treat or ameliorate a medical condition in e.g.  
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility to  
 CC a pathological condition. The antibodies to the polypeptides can also be  
 CC used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. The present sequence  
 CC represents the amino acid sequence of human heparanase-like protein,  
 CC splice variant #1  
 XX  
 SQ Sequence 592 AA;  
 QY Query Match 99.7%; Score 3078; DB 4; Length 592;  
 DB Best Local Similarity 99.7%; Pred. No. 1,3e-303;  
 Matches 590; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSQAGRRPYPVRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALLLHLSLSQAGRRPYPVRAAGLKEKTL 60  
 QY 61 ILDDVSTKNPVRTVNEFLSLQDPSIIHDGMDFLSSKRLVTLARGLSPAFIRFGSKRT 120  
 DB 61 ILDDVSTKNPVRTVNEFLSLQDPSIIHDGMDFLSSKRLVTLARGLSPAFIRFGSKRT 120  
 QY 121 DFLQFQNLRNPAKRGGRGPPYLYKXVEDDVIYSDVALDKOKGCKIAQHPDWLVLOREK 180  
 DB 121 DFLQFQNLRNPAKRGGRGPPYLYKXVEDDVIYSDVALDKOKGCKIAQHPDWLVLOREK 180  
 QY 121 DFLQFQNLRNPAKRGGRGPPYLYKXVEDDVIYSDVALDKOKGCKIAQHPDWLVLOREK 180  
 DB 121 DFLQFQNLRNPAKRGGRGPPYLYKXVEDDVIYSDVALDKOKGCKIAQHPDWLVLOREK 180  
 QY 181 AAQMHLYLLKEQFNTYSNLLITARSIDKLYNPAQCSGLHLIPALNALRPNPNNSNSS 240  
 DB 181 AAQMHLYLLKEQFNTYSNLLITARSIDKLYNPAQCSGLHLIPALNALRPNPNNSNSS 240  
 QY 241 ALSLIKTSASAKKINISWELGNEPNNTMGRANVSGQLGQDYIQLKSLQPLIYSRAS 300  
 DB 241 ALSLIKTSASAKKINISWELGNEPNNTMGRANVSGQLGQDYIQLKSLQPLIYSRAS 300  
 QY 241 ALSLIKTSASAKKINISWELGNEPNNTMGRANVSGQLGQDYIQLKSLQPLIYSRAS 300  
 DB 241 ALSLIKTSASAKKINISWELGNEPNNTMGRANVSGQLGQDYIQLKSLQPLIYSRAS 300  
 QY 301 LYGPNIQPRKXVYALADGFMKVGASTVDVAVTMOHCYIDGRVYKMDFLKTRLADTLSDQ 360  
 DB 301 LYGPNIQPRKXVYALADGFMKVGASTVDVAVTMOHCYIDGRVYKMDFLKTRLADTLSDQ 360  
 QY 301 LYGPNIQPRKXVYALADGFMKVGASTVDVAVTMOHCYIDGRVYKMDFLKTRLADTLSDQ 360  
 DB 301 LYGPNIQPRKXVYALADGFMKVGASTVDVAVTMOHCYIDGRVYKMDFLKTRLADTLSDQ 360  
 QY 361 IRKIQKVNTYTPGKKIWLBEVVTTSAGTNNLSDSYAGTFLMNTGMLANQSIDVIVIR 420  
 DB 361 IRKIQKVNTYTPGKKIWLBEVVTTSAGTNNLSDSYAGTFLMNTGMLANQSIDVIVIR 420  
 QY 361 IRKIQKVNTYTPGKKIWLBEVVTTSAGTNNLSDSYAGTFLMNTGMLANQSIDVIVIR 420  
 DB 361 IRKIQKVNTYTPGKKIWLBEVVTTSAGTNNLSDSYAGTFLMNTGMLANQSIDVIVIR 420  
 QY 421 HSPFHGYNHLYDONENPLPDYWLSLYKRLIGPKYLAHVAGLORRPPRGRVIRDKLRI 480  
 DB 421 HSPFHGYNHLYDONENPLPDYWLSLYKRLIGPKYLAHVAGLORRPPRGRVIRDKLRI 480  
 QY 421 HSPFHGYNHLYDONENPLPDYWLSLYKRLIGPKYLAHVAGLORRPPRGRVIRDKLRI 480  
 DB 421 HSPFHGYNHLYDONENPLPDYWLSLYKRLIGPKYLAHVAGLORRPPRGRVIRDKLRI 480  
 QY 481 YACCTNHNHNHYVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGELKSKSV 540  
 DB 481 YACCTNHNHNHYVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGELKSKSV 540  
 QY 481 YACCTNHNHNHYVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGELKSKSV 540  
 DB 481 YACCTNHNHNHYVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYQEGELKSKSV 540  
 QY 541 QUNGQPLVWVDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592  
 DB 541 QUNGQPLVWVDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592  
 RESULT 4  
 ID AAB85215 standard; protein; 592 AA.  
 XX AAB85215;  
 AC  
 XX 07-SEP-2001 (first entry)  
 DT  
 XX  
 DE Heparanase-like protein Hpa2 splice variant #1.  
 XX  
 KM Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KM cytostatic; neuroprotective; cerebroprotective; immunosuppressive;  
 KM antidiabetic; antiarthritic; antiinflammatory; antilastematic;  
 KM antidiabetic; antiarteriosclerotic; vulnerary.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 237

FT /label= unknown  
 FT /note="encoded by ANC"  
 XX  
 XX MO200146392-A2.  
 XX  
 XX 28-JUN-2001.  
 XX  
 XX 21-DEC-2000; 2000MO-GB004963.  
 XX  
 XX 22-DEC-1999; 99GB-00030392.  
 PR 07-APR-2000; 2000GB-00008713.  
 XX  
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX  
 XX McKenzie EA, Stamps AC, Terrett JA, Tyson KL;  
 XX  
 XX WPI; 2001-418056/44.  
 DR N-PSDB; AAE22671.  
 XX  
 XX Novel homologs of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 XX  
 XX Claim 1; Fig 1; 97pp; English.  
 PS  
 XX The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and inflammation. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the  
 CC amino acid sequence of the largest splice variant of the heparanase-like  
 CC protein Hpa2 of the invention  
 XX  
 XX Sequence 592 AA;  
 SQ  
 Query March 99.4%; Score 3071; DB 4; Length 592;  
 Best Local Similarity 99.5%; Pred. No. 6.9e-303;  
 Matches 589; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MRVLCAPEAMPSSNSRPPACIAPGALYLLALHTLSISQAGDRRPLVDRAAGLKEKTL 60  
 DB 1 MRVLCAPEAMPSSNSRPPACIAPGALYLLALHTLSISQAGDRRPLVDRAAGLKEKTL 60  
 QY 61 ILIDVSTKPNRTNENPLSLQDPSITIHGMDLDFSSKRLVLTARLSAPFRFGKRT 120  
 DB 61 ILIDVSTKPNRTNENPLSLQDPSITIHGMDLDFSSKRLVLTARLSAPFRFGKRT 120  
 QY 121 DFLQFQNLNRPFAKSGRGPDPYLYKNYEDDIVRSVALDQKQCKIAQHDPVMLVLOREK 180  
 DB 121 DFLQFQNLNRPFAKSGRGPDPYLYKNYEDDIVRSVALDQKQCKIAQHDPVMLVLOREK 180  
 QY 121 DFLQFQNLNRPFAKSGRGPDPYLYKNYEDDIVRSVALDQKQCKIAQHDPVMLVLOREK 180  
 DB 121 DFLQFQNLNRPFAKSGRGPDPYLYKNYEDDIVRSVALDQKQCKIAQHDPVMLVLOREK 180  
 QY 181 AAQNHVLLKQESNTTNSLILTRASLDKYNPDCGGLHIFMLNLRPNPNNSWSSS 240  
 DB 181 AAQNHVLLKQESNTTNSLILTRASLDKYNPDCGGLHIFMLNLRPNPNNSWSSS 240  
 QY 181 AAQNHVLLKQESNTTNSLILTRASLDKYNPDCGGLHIFMLNLRPNPNNSWSSS 240  
 DB 181 AAQNHVLLKQESNTTNSLILTRASLDKYNPDCGGLHIFMLNLRPNPNNSWSSS 240  
 QY 241 ALSLIKTSASKKYINISWELGNEPNRYTHMGRANVSGQDGYQLKSLQPIRIYSRAS 300  
 DB 241 ALSLIKTSASKKYINISWELGNEPNRYTHMGRANVSGQDGYQLKSLQPIRIYSRAS 300  
 QY 301 LYGNIGRPRKNVALLDGFMKVAGSTVDANTWQCHTIDGVYKVMPLKTRLLDTSDQ 360  
 DB 301 LYGNIGRPRKNVALLDGFMKVAGSTVDANTWQCHTIDGVYKVMPLKTRLLDTSDQ 360

QY 361 IRKIQKVNVTYTPGKKIWEIGVTTSAAGTNNISDSYAGFLMLNTLIGMIANQGDIVVIR 420  
 DB 361 IRKIQKVNVTYTPGKKIWEIGVTTSAAGTNNISDSYAGFLMLNTLIGMIANQGDIVVIR 420  
 QY 421 HSPFDHGYNHLVQONENPLPDYMLSLYKRLIDPKYLAHVAGLQRRPRGRVIRPKLRI 480  
 DB 421 HSPFDHGYNHLVQONENPLPDYMLSLYKRLIDPKYLAHVAGLQRRPRGRVIRPKLRI 480  
 QY 481 YAHCTNNHNNHYVGSITLFIINLRGRKKIKLAGLRDCLVHQYLLQPYGQGLSKSKSV 540  
 DB 481 YAHCTNNHNNHYVGSITLFIINLRGRKKIKLAGLRDCLVHQYLLQPYGQGLSKSKSV 540  
 QY 541 QINQGPLVMDDGTLPELKRPLRAGTLYIPPTWGFYVKNVNLACRYR 592  
 DB 541 QINQGPLVMDDGTLPELKRPLRAGTLYIPPTWGFYVKNVNLACRYR 592  
 RESULT 5  
 AAE18326  
 ID AAE18326 standard; protein; 582 AA.  
 XX  
 AC AAE18326;  
 XX  
 XX 07-MAY-2002 (first entry)  
 DT  
 XX  
 DE Human heparanase-2AB splice variant protein.  
 XX  
 XX Human; heparanase-2AB; Hep-2; wound healing; angiogenesis; restenosis;  
 XX atherosclerosis; neurodegenerative disease; inflammation; prostate;  
 XX viral infection; autoimmune lesion; renal failure; pancreatic cancer;  
 XX dystrophic muscular disease; heart disease; gene therapy; enzyme.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200204645-A2.  
 XX  
 XX 17-JAN-2002.  
 XX  
 XX 12-JUL-2001; 2001WO-BP008094.  
 PF  
 XX 12-JUL-2000; 2000EP-00202442.  
 PR  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 XX  
 XX David G, Duerr J;  
 PI  
 XX WPI; 2002-171719/22.  
 DR N-PSDB; AAD29202.  
 XX  
 XX Heparanase-2 polypeptides and polynucleotides, useful for useful in wound  
 PT healing, angiogenesis, and for treating restenosis, atherosclerosis,  
 PT inflammation, neurodegenerative diseases, and viral infections.  
 XX  
 XX Claim 1; Page 38-40; 54pp; English.  
 PS  
 XX The invention relates to human heparanase-2 (Hep-2) polypeptides and  
 CC polynucleotides. Heparanase-2 protein is useful in wound healing,  
 CC angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases,  
 CC inflammation and viral infections, as well as in neutralising plasma  
 CC heparin as a potential replacement of protamine. Antiheparanase-2  
 CC antibodies may be used for immunodetection and diagnosis of  
 CC micrometastases, autoimmune lesions, renal failure in biopsy specimens,  
 CC plasma samples and body fluids. Molecules, which can agonise or  
 CC antagonise heparanase 2 catalytic activity may also be used as a  
 CC medicament. Polymorphisms in the polynucleotide sequence are useful in  
 CC the identification of individuals having a predisposition to acquire  
 CC diseases resulting from an increased or decreased expression of their  
 CC activity. Such molecules can be used to treat pancreatic cancer,  
 CC dystrophic muscular diseases and or heart diseases. Polynucleotides of  
 CC the invention are used in gene therapy. The present sequence is human  
 CC heparanase-2AB splice variant protein  
 XX  
 XX Sequence 582 AA;  
 SQ

Query Match 97.4%; Score 3007; DB 5; Length 582;  
 Best Local Similarity 99.1%; Pred. No. 2.2e-296;  
 Matches 577; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 11 MSSNSRPPACAPGALYLLALHLHLSLSSQAGDRRPLPVDRAAGLKEKTLILDVSTNPF 70  
 1 MSSNSRPPACAPGALYLLALHLHLSLSSQAGDRRPLPVDRAAGLKEKTLILDVSTNPF 60  
 DB 1 MSSNSRPPACAPGALYLLALHLHLSLSSQAGDRRPLPVDRAAGLKEKTLILDVSTNPF 60  
 QY 71 VRTVNEFLSLQDPSIHDGMDFLSSKRLVTLARGSLPAFLRFGKRTDPLQONTN 130  
 DB 61 VRTVNEFLSLQDPSIHDGMDFLSSKRLVTLARGSLPAFLRFGKRTDPLQONTN 120  
 QY 131 PAKSRGGPDDYLLKRYEDDIVRSVVALDKQCKIAQHPDVMVLQREKAQOMLVLLK 190  
 DB 121 PAKSRGGPDDYLLKRYEDDIVRSVVALDKQCKIAQHPDVMVLQREKAQOMLVLLK 180  
 QY 191 EGFNTYSMLILTARSLDKLYNFADCSGLHLIFALNALRRNPNNSWSSALSILKYSAS 250  
 DB 181 EGFNTYSMLILTARSLDKLYNFADCSGLHLIFALNALRRNPNNSWSSALSILKYSAS 240  
 QY 251 KKNINISWELGNEPNNTYMGRAVNSQLGKDYIOUKSLQPIRYSRASLYGPNIGRPR 310  
 DB 241 KKNINISWELGNEPNNTYMGRAVNSQLGKDYIOUKSLQPIRYSRASLYGPNIGRPR 300  
 QY 311 KNYIALLDGFMKAGSTVDVAVTWQHCYIDGRVVKVMDPLKTRLLDPLSDQIRIKQVNT 370  
 DB 301 KNYIALLDGFMKAGSTVDVAVTWQHCYIDGRVVKVMDPLKTRLLDPLSDQIRIKQVNT 360  
 QY 371 YTFGKKIMLEGVTTTSAAGTNNISDSYAAGFLMNTLGMLANOGIDVYIRHSFPDHGYNH 430  
 DB 361 YTFGKKIMLEGVTTTSAAGTNNISDSYAAGFLMNTLGMLANOGIDVYIRHSFPDHGYNH 420  
 QY 431 LVQONTNPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRYDKLRIYAHCTNNHH 490  
 DB 421 LVQONTNPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRYDKLRIYAHCTNNHH 480  
 QY 491 NYVRSITFTFIINLHRSRKKIKLAGLRDLVHOYLLQPYGQGLKSKSVQANGPLVNV 550  
 DB 481 NYVRSITFTFIINLHRSRKKIKLAGLRDLVHOYLLQPYGQGLKSKSVQANGPLVNV 540  
 QY 551 DDGTLBELKRPPLRAGRTLVIPVTMGFFVYKVNALLACRYR 592  
 DB 541 DDGTLBELKRPPLRAGRTLVIPVTMGFFVYKVNALLACRYR 582

RESULT 6  
 ID AAY97633 standard; protein; 538 AA.  
 AC AAY97633;  
 DT 20-APR-2001 (first entry)  
 DE Human heparanase, hnhp1 p99 form, protein sequence.  
 XX Heparanase; hnhp1; wound healing; angiogenesis; restenosis; Scrape;  
 KM atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;  
 KM neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;  
 KM gene therapy; human.  
 OS Homo sapiens.  
 XX  
 XX  
 FT Key Location/Qualifiers  
 FT Misc-difference 305 /note="encoded by GAC"  
 XX  
 XX WO200100643-A2.  
 XX 04-JAN-2001.  
 XX 19-JUN-2000; 2000MO-IL000358.  
 XX

PR 25-JUN-1999; 99US-0140801P.  
 XX (INST-) INSIGHT STRATEGY & MARKETING LTD.  
 XX  
 XX Pecker I, Michael I, Itzhaki H;  
 XX  
 XX WPI; 2001-137930/14.  
 DR N-PSDB; AAA91098.  
 XX  
 PT New polynucleotides and polypeptides that are distantly homologous to  
 PT heparanase, useful in wound healing, as well as in gene therapy protocols  
 PT for angiogenesis, restenosis, atherosclerosis, or inflammation.  
 XX  
 PS Claim 10; Page 61-62; 67pp; English.  
 CC This sequence represents a heparanase of the invention. The heparanase  
 CC DNA and protein sequences are useful in wound healing, angiogenesis,  
 CC restenosis, atherosclerosis, inflammation, pulmonary disease, and  
 CC neurodegenerative diseases (such as Scrape, Alzheimer's disease, and  
 CC Creutzfeldt-Jakob disease) or viral infections. The heparanase coding  
 CC sequence is particularly useful in gene therapy  
 XX  
 XX Sequence 538 AA:

Query Match 89.7%; Score 2770; DB 4; Length 538;  
 Best Local Similarity 90.5%; Pred. No. 2.6e-272;  
 Matches 536; Conservative 1; Mismatches 1; Indels 54; Gaps 1;

QY 1 MRVLCAPPEAMPSSNSRPPACAPGALYLLALHLHLSLSSQAGDRRPLPVDRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSSNSRPPACAPGALYLLALHLHLSLSSQAGDRRPLPVDRAAGLKEKTL 60  
 QY 61 ILDVDSTKNPRTVNEFLSLQDPSIHDGMDFLSSKRLVTLARGSLPAFLRFGKRT 120  
 DB 61 ILDVDSTKNPRTVNEFLSLQDPSIHDGMDFLSSKRLVTLARGSLPAFLRFGKRT 120  
 QY 121 DFLQONTNPNKARGGPGDYVLYKXVEDDIVASDVALLDKQCKIAQHPDVMVLQREK 180  
 DB 121 DFLQONTNPNKARGGPGDYVLYKXVEDDIVASDVALLDKQCKIAQHPDVMVLQREK 149  
 QY 181 AAQMHVLLKEQFENTYSMLILTARSLDKLYNFADCSGLHLIFALNALRRNPNNSWSS 240  
 DB 150 -----ARSLDKLYNFADCSGLHLIFALNALRRNPNNSWSS 186  
 QY 241 ALSILKYSASAKKNIISWELGNEPNNTYMGRAVNSQLGKDYIOUKSLQPIRYSRAS 300  
 DB 187 ALSILKYSASAKKNIISWELGNEPNNTYMGRAVNSQLGKDYIOUKSLQPIRYSRAS 246  
 QY 301 LYGPNIQRPKNVIALLDGFMKAGSTVDVAVTWQHCYIDGRVVKVMDPLKTRLLDPLSDQ 360  
 DB 247 LYGPNIQRPKNVIALLDGFMKAGSTVDVAVTWQHCYIDGRVVKVMDPLKTRLLDPLSDQ 306  
 QY 361 IRKIQKVVNTYTPGKCTWLBGVTTTSAAGTNNISDSYAAGFLMNTLGMLANOGIDVYIR 420  
 DB 307 IRKIQKVVNTYTPGKCTWLBGVTTTSAAGTNNISDSYAAGFLMNTLGMLANOGIDVYIR 366  
 QY 421 HSFPDHGYNHLYVDONFPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRYDKLRI 480  
 DB 367 HSFPDHGYNHLYVDONFPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRYDKLRI 426  
 QY 481 YAHCTNNHNNYVRSITFTFIINLHRSRKKIKLAGLRDLVHOYLLQPYGQGLKSKSV 540  
 DB 427 YAHCTNNHNNYVRSITFTFIINLHRSRKKIKLAGLRDLVHOYLLQPYGQGLKSKSV 486  
 QY 541 QLNGQPLVMVDGTLPELKRPPLRAGRTLVIPVTMGFFVYKVNALLACRYR 592  
 DB 487 QLNGQPLVMVDGTLPELKRPPLRAGRTLVIPVTMGFFVYKVNALLACRYR 538

RESULT 7  
 ID AAB85216 standard; protein; 534 AA.  
 XX  
 XX AAB85216  
 XX 19-JUN-2000; 2000MO-IL000358.  
 XX

AC AAB85216;  
 XX 07-SEP-2001 (first entry)  
 DE Heparanase-like protein Hpa2 splice variant #2.  
 XX Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;  
 KM cytosolic; neuroprotective; cerebroprotective; immunosuppressive;  
 KM antiproliferative; nototropic; antiinflammatory; antiarthritic; antidiabetic;  
 KM antidiabetic; antiarteriosclerotic; vulnery.  
 OS Homo sapiens.  
 XX MO200146392-A2.  
 XX 28-JUN-2001.  
 PD 21-DEC-2000; 2000WO-GB004963.  
 PF 22-DEC-1999; 99GB-000030392.  
 PR 07-APR-2000; 2000GB-00008713.  
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PA McKenzie EA, Stamps AC, Terrett JA, Tyson XL;  
 PI MPI; 2001-418056/44.  
 XX N-PSDB; AAB22672.  
 DR Novel homologs of heparanase, present in three splice variants, useful  
 PT for identifying agents that modulate heparanase, useful in the treatment  
 PT and/or prophylaxis of abnormal levels of heparanase.  
 XX Claim 1; Fig 2; 97JP; English.  
 XX The invention provides a homologue to heparanase which is present in  
 CC three splice variants. The heparanase homologue polypeptide is useful in  
 CC the treatment of a human or non-human animal or for use in diagnosis.  
 CC Vectors comprising the heparanase homologue polynucleotides are useful in  
 CC the transformation or transfection of a prokaryotic or eukaryotic host.  
 CC The modulators of the polypeptide are useful in the manufacture of a  
 CC medicament for the treatment and/or prophylaxis of a condition/disease  
 CC associated with abnormal levels of the heparanase homologue, including  
 CC cancer, central nervous system (CNS) and neurodegenerative diseases,  
 CC cardiovascular diseases such as restenosis following angioplasty and  
 CC atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,  
 CC allografts, inflammatory diseases, arthritis, vascular restenosis, tumour  
 CC growth and progression, asthma, Alzheimer's disease, diabetic  
 CC retinopathy, wound healing and research. The polypeptide is also  
 CC useful in diagnosis and research. The present sequence represents the  
 CC amino acid sequence of the mid-sized splice variant of the heparanase-  
 CC like protein Hpa2 of the invention  
 XX Sequence 534 AA;  
 SO  
 Query March 88.6%; Score 2736; DB 4; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 7.5e-269; Indels 58; Gaps 1;  
 Matches 534; Conservative 1; Mismatches 1;  
 QY 1 MRVLCAPPEAMPSSNSRPPACLAGALYLLALHLISISQAGRRPLVDRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSSNSRPPACLAGALYLLALHLISISQAGRRPLVDRAAGLKEKTL 60  
 QY 1 IILDVSTKNTVNTNEMFLSLQDPSITIHGMDPLSSKRLVTLAAGLSPAFRFGGKRT 120  
 DB 61 IILDVSTKNTVNTNEMFLSLQDPSITIHGMDPLSSKRLVTLAAGLSPAFRFGGKRT 120  
 QY 61 IILDVSTKNTVNTNEMFLSLQDPSITIHGMDPLSSKRLVTLAAGLSPAFRFGGKRT 120  
 DB 121 DFLQFQMLRPAKSGGPGPDYLLKNYEDDIVSDVALDKQKCKIAQHPDVMLEQREX 180  
 QY 121 DFLQFQMLRPAKSGGPGPDYLLKNYEDDIVSDVALDKQKCKIAQHPDVMLEQREX 180  
 DB 121 DFLQFQMLRPAKSGGPGPDYLLKNYEDDIVSDVALDKQKCKIAQHPDVMLEQREX 180  
 QY 181 AAGMHLVLKEQPSNTYSNLTLT----- 203  
 DB 181 AAGMHLVLKEQPSNTYSNLTLT----- 203

DB 181 AAGMHLVLKEQPSNTYSNLTLT----- 203  
 QY 241 ALSLLKYSASKKYNISWEIGNEPNRTMGRANVSGQKDYIOLKSLQIRIYSRAS 300  
 DB 204 -----EPNNYRTMGRANVSGQKDYIOLKSLQIRIYSRAS 242  
 QY 301 LYGPNIGRPKKNVIALLDGFMKAGSTVDVAVTMOHCYIDGRVVKWDFLKTBLDLSQ 360  
 DB 243 LYGPNIGRPKKNVIALLDGFMKAGSTVDVAVTMOHCYIDGRVVKWDFLKTBLDLSQ 302  
 QY 361 IRKIQKVVNTYTPGKKIMLEGGVTTTSGAGTNNLSDSYAGFLMNTLGMLANQSIDVVR 420  
 DB 303 IRKIQKVVNTYTPGKKIMLEGGVTTTSGAGTNNLSDSYAGFLMNTLGMLANQSIDVVR 362  
 QY 421 HSPFDGYNHLVDQNTNPLPDYVLSLYRKLIGPKYLAHVAGLQKRPGRGVIRDKLRI 480  
 DB 363 HSPFDGYNHLVDQNTNPLPDYVLSLYRKLIGPKYLAHVAGLQKRPGRGVIRDKLRI 422  
 QY 481 YACHTNHNHNHYVRSITLFIINLHRSKKIKLAGTLRDKLVHGYLLQFYGOGLKSKSV 540  
 DB 423 YACHTNHNHNHYVRSITLFIINLHRSKKIKLAGTLRDKLVHGYLLQFYGOGLKSKSV 482  
 QY 541 QLNQGPLVWDDGTLPELKPRLAAGRTLVIPVPMGFYVVKVNALACRYR 592  
 DB 483 QLNQGPLVWDDGTLPELKPRLAAGRTLVIPVPMGFYVVKVNALACRYR 534  
 RESULT 8  
 ID ABE69310 standard; protein; 534 AA.  
 AC ABE69310;  
 XX 20-JAN-2003 (first entry)  
 DE Human polypeptide SEQ ID NO 1357.  
 XX Human, genome mapping; gene therapy; food supplement; virus; fungus;  
 KM cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KM Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KM arthritis; cytotoxic; immunomodulator; nototropic; neuroprotective;  
 KM antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KM haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KM antidiabetic.  
 OS Homo sapiens.  
 XX WO200270539-A2.  
 XX 12-SEP-2002.  
 PD 05-MAR-2002; 2002WO-US005095.  
 PF 05-MAR-2001; 2001US-00799451.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
 PI Wehman T, Wang J, Wang D, Drmanac RT;  
 DR MPI; 2002-759812/82.  
 DR N-PSDB; ABE21527.  
 XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.  
 PS Claim 9; SEQ ID NO 1357; 1012pp + Sequence Listing; English.  
 CC The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences (AB211119-  
 CC AB212066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, liver  
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
 CC arthritis, etc. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 534 AA:

Query Match 88.6%; Score 2736; DB 5; Length 534;

Best Local Similarity 89.9%; Pred. No. 7.5e-269;

Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

QY 1 MRVLCAPPEAMPSPNSRPPACLPALYALLLHLTSLSSQAGRRLPVDRAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSPNSRPPACLPALYALLLHLTSLSSQAGRRLPVDRAAGLKEKTL 60  
 QY 61 ILDVDSTKRNVTNENFSLQDPSIHDGMLDFTSSKRLVTLARGLSPAFRFGGKRT 120  
 DB 61 ILDVDSTKRNVTNENFSLQDPSIHDGMLDFTSSKRLVTLARGLSPAFRFGGKRT 120  
 QY 121 DFLQFQNLNRPASRGSGPDDYIKNYEDDIVRSVVALDKQCKIAQHPDVLVTOREK 180  
 DB 121 DFLQFQNLNRPASRGSGPDDYIKNYEDDIVRSVVALDKQCKIAQHPDVLVTOREK 180  
 QY 181 AAOQHVLTLKEOPSNTSNLITPARSLDKYFPADSGHLIPALNALRPNMWSSS 240  
 DB 181 AAOQHVLTLKEOPSNTSNLITPARSLDKYFPADSGHLIPALNALRPNMWSSS 240  
 QY 241 ALSLIKYSASKKNIISWELNEPNRYRTMGRAVNSQLGKDYIQLKSLIPIRISRAS 300  
 DB 241 ALSLIKYSASKKNIISWELNEPNRYRTMGRAVNSQLGKDYIQLKSLIPIRISRAS 300  
 QY 204 -----EPNNYRTMGRAVNSQLGKDYIQLKSLIPIRISRAS 242  
 DB 204 -----EPNNYRTMGRAVNSQLGKDYIQLKSLIPIRISRAS 242  
 QY 301 LYGNIRPRKNVYALLDGPMKVGSTVDVWMOHCYIDGRVYKVMDFELRLDLTSDQ 360  
 DB 301 LYGNIRPRKNVYALLDGPMKVGSTVDVWMOHCYIDGRVYKVMDFELRLDLTSDQ 360  
 QY 243 LYGNIRPRKNVYALLDGPMKVGSTVDVWMOHCYIDGRVYKVMDFELRLDLTSDQ 302  
 DB 243 LYGNIRPRKNVYALLDGPMKVGSTVDVWMOHCYIDGRVYKVMDFELRLDLTSDQ 302  
 QY 361 IRKIQKVNNTYTPGKTIWLEGVVTTSAGTNNLSDSYAAQFLMNTLGMLANOGIDVIR 420  
 DB 361 IRKIQKVNNTYTPGKTIWLEGVVTTSAGTNNLSDSYAAQFLMNTLGMLANOGIDVIR 420  
 QY 421 HSFDHGYNHLVDGNFPLPRYMSLYKRLIGKYLAVVAGQRPREGRIYRDLRI 480  
 DB 421 HSFDHGYNHLVDGNFPLPRYMSLYKRLIGKYLAVVAGQRPREGRIYRDLRI 480  
 QY 363 HSFDHGYNHLVDGNFPLPRYMSLYKRLIGKYLAVVAGQRPREGRIYRDLRI 422  
 DB 363 HSFDHGYNHLVDGNFPLPRYMSLYKRLIGKYLAVVAGQRPREGRIYRDLRI 422  
 QY 481 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTLRDLKVOYLLQPYGOBGLSKSV 540  
 DB 481 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTLRDLKVOYLLQPYGOBGLSKSV 540  
 QY 423 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTLRDLKVOYLLQPYGOBGLSKSV 482  
 DB 423 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTLRDLKVOYLLQPYGOBGLSKSV 482  
 QY 541 QLNQGPLWVDGTLPELKPRPLAAGRTLVPTTGMGFYVKNVNAACRYR 592  
 DB 541 QLNQGPLWVDGTLPELKPRPLAAGRTLVPTTGMGFYVKNVNAACRYR 592  
 QY 483 QLNQGPLWVDGTLPELKPRPLAAGRTLVPTTGMGFYVKNVNAACRYR 534  
 DB 483 QLNQGPLWVDGTLPELKPRPLAAGRTLVPTTGMGFYVKNVNAACRYR 534

RESULT 9

AAMS0337 standard; protein, 534 AA.

AC AAMS0337;

DT 04-FEB-2002 (first entry)

DE Human prepro-heparanase II.

KW Heparanase II; human; vulnery; angiogenesis inhibitor;  
 KW antiinflammatory; cytostatic; therapy; diagnosis.

OS Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..41

XX Protein 42..534

XX Protein 42..129

XX Protein 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129

XX Modified-site 42..129



CC treating or preventing metastasis, cancer, CNS and neurodegenerative  
 CC diseases, inflammation and cardiovascular disease such as restenosis  
 CC following angioplasty and atherosclerosis. Heparanase II is useful for  
 CC accelerating wound healing, blocking angiogenesis, degradation of heparin  
 CC and neutralization of heparin's anticoagulant properties during surgery.  
 CC Heparanase or an agent that enhances heparanase activity can also be  
 CC infused into the vasculature to block accumulation and diapedesis of  
 CC neutrophils at sites of inflammation with or without added domains to  
 CC confer selectivity in delivery

XX Sequence 534 AA.

Query Match 88.6%; Score 2736; DB 5; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 7,5e-269;  
 Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

QY 1 MRVLCAPPEAMPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDAAGLKEKTL 60  
 DB 1 MRVLCAPPEAMPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDAAGLKEKTL 60  
 QY 61 ILDDSTKPNFRTVNEFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFGSKRT 120  
 DB 61 ILDDSTKPNFRTVNEFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFGSKRT 120  
 QY 121 DFIQFQWLRPAKSRGPGPDYLLKNEYDYASDVALKQKCKIAQHPDVMVLQREK 180  
 DB 121 DFIQFQWLRPAKSRGPGPDYLLKNEYDYASDVALKQKCKIAQHPDVMVLQREK 180  
 QY 181 AAGMHLVLEKQPSNTSYNLIILFARSLDKYNFADCSGLILFALNALRRPNNSWSSS 240  
 DB 181 AAGMHLVLEKQPSNTSYNLIILFARSLDKYNFADCSGLILFALNALRRPNNSWSSS 240  
 QY 241 ALSLLKYSASKYNNISWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRAS 300  
 DB 241 ALSLLKYSASKYNNISWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRAS 300  
 QY 301 LYGPNIIGRPKNVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFKTRLDLTLSDQ 360  
 DB 301 LYGPNIIGRPKNVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFKTRLDLTLSDQ 360  
 QY 361 IRRIQKRVNTYTPGKKIMLEGVVTTSGAGTNNLSDSYAAGFLMNTLGMANQSIDVIR 420  
 DB 361 IRRIQKRVNTYTPGKKIMLEGVVTTSGAGTNNLSDSYAAGFLMNTLGMANQSIDVIR 420  
 QY 421 HSPFDGYNHLVQONRPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRPKLRI 480  
 DB 421 HSPFDGYNHLVQONRPLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRPKLRI 480  
 QY 481 YACHTNHNHNNYVRSITLFTIIMHRSRKKIKLAGTLRDLVHQVYLLQPYGQGLSKSV 540  
 DB 481 YACHTNHNHNNYVRSITLFTIIMHRSRKKIKLAGTLRDLVHQVYLLQPYGQGLSKSV 540  
 QY 541 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRYR 592  
 DB 541 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRYR 592  
 QY 483 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRYR 534  
 DB 483 QLNGQPLVMVDGTLPELKRPRLRAGTTLVTPYTMGFYVKNVNLACRYR 534

RESULT 10  
 AAE18327  
 ID AAE18327 standard; protein; 528 AA.  
 XX  
 AC AAE18327;  
 XX  
 DT 07-MAY-2002 (first entry)  
 XX  
 DE Human heparanase-2A splice variant protein.  
 XX  
 KW Human, heparanase-2A; Hep-2; wound healing; angiogenesis; restenosis;  
 KW atherosclerosis; neurodegenerative disease; inflammation; procarcane;  
 KW viral infection; autoimmune lesion; renal failure; pancreatic cancer;  
 KW dystrophic muscular disease; heart disease; gene therapy; enzyme.  
 XX  
 OS Homo sapiens.

XX MO200204645-A2.  
 XX 17-JAN-2002.  
 XX 12-JUL-2001; 2001WO-EP008094.  
 XX 12-JUL-2000; 2000EP-00202442.  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 XX David G. Duert J;  
 XX WPI; 2002-171719/22.  
 XX N-PSDB; RAD29204.

Heparanase-2 polypeptides and polymucleotides, useful for useful in wound healing, angiogenesis, and for treating restenosis, atherosclerosis, inflammation, neurodegenerative diseases, and viral infections.

Disclosure; Page 45-46; 54pp; English.

The invention relates to human heparanase-2 (Hep-2) polypeptides and polymucleotides. Heparanase-2 protein is useful in wound healing, angiogenesis, restenosis, atherosclerosis, neurodegenerative diseases, inflammation and viral infections, as well as in neutralizing plasma heparin as a potential replacement of protamine. Antiheparanase-2 antibodies may be used for immunodetection and diagnosis of micrometastases, autoimmune lesions, renal failure in biopsy specimens, plasma samples and body fluids. Molecules, which can agonize or antagonize heparanase 2 catalytic activity may also be used as a medicament. Polymorphisms in the polymucleotide sequence are useful in the identification of individuals having a predisposition to acquire diseases resulting from a increased or decreased expression of their activity. Such molecules can be used to treat pancreatic cancer, CC dystrophic muscular diseases and or heart diseases. Polymucleotides of the invention are used in gene therapy. The present sequence is human heparanase-2A splice variant protein

XX Sequence 528 AA;

Query Match 88.1%; Score 2720; DB 5; Length 528;  
 Best Local Similarity 90.4%; Pred. No. 3.1e-267;  
 Matches 526; Conservative 1; Mismatches 1; Indels 54; Gaps 1;

QY 11 MPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDAAGLKEKTLILDDSTKPN 70  
 DB 1 MPSSNSRPACLAGALYALILHLISLSQAGRRPLPYDAAGLKEKTLILDDSTKPN 70  
 QY 71 VRTNENFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFGSKRTDFOQLNRN 130  
 DB 71 VRTNENFSLQDPSIIHDGMLDPLSKRLVTLARGLSPAFRFGSKRTDFOQLNRN 130  
 QY 131 PAKSRGPGPDYLLKNEYDYASDVALKQKCKIAQHPDVMVLQREKAAQMHVLVLK 190  
 DB 131 PAKSRGPGPDYLLKNEYDYASDVALKQKCKIAQHPDVMVLQREKAAQMHVLVLK 190  
 QY 191 EGFSTYENLITLARSIDKYNFADCSGLILFALNALRRPNNSWSSSALSLKYSAS 250  
 DB 191 EGFSTYENLITLARSIDKYNFADCSGLILFALNALRRPNNSWSSSALSLKYSAS 250  
 QY 251 KKYNIWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRASLYGPNIRPR 310  
 DB 251 KKYNIWELGNEPNNTYRMHGRAVNSQIGKDYTLQKSLLOPIRYSRASLYGPNIRPR 310  
 QY 311 KNVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFKTRLDLTLSDQIRKIQKVT 370  
 DB 311 KNVIALIDGFMKVAAGSTVAVTWQHCYIDGRVYKWMDFKTRLDLTLSDQIRKIQKVT 370  
 QY 371 YTPGKKIMLEGVVTTSGAGTNNLSDSYAAGFLMNTLGMANQSIDVIRHSPFDGYNH 430  
 DB 371 YTPGKKIMLEGVVTTSGAGTNNLSDSYAAGFLMNTLGMANQSIDVIRHSPFDGYNH 430

QY 431 LVONENPLDPYMLSLIKKELIGPKVLAVHAGLQKRPGRVIRDPKLRITAHCTNHHN 490  
 DB 367 LVONENPLDPYMLSLIKKELIGPKVLAVHAGLQKRPGRVIRDPKLRITAHCTNHHN 426  
 QY 491 NYVRSITLFIINLHRSRKKIKLAGTLRDLVHQLQPYGQGLSKSVQVNGQPLVWV 550  
 DB 427 NYVRSITLFIINLHRSRKKIKLAGTLRDLVHQLQPYGQGLSKSVQVNGQPLVWV 486  
 QY 551 DDGTLPELKRPLRAGRTVIRPPTMGFPYVKNVNNALACRYR 592  
 DB 487 DDGTLPELKRPLRAGRTVIRPPTMGFPYVKNVNNALACRYR 528

## RESULT 11

AAB84664 standard; protein; 492 AA.

AC AAB84664;

XX 05-SEP-2001 (first entry)

DE Amino acid sequence of human heparanase-like polypeptide.

KW Human; heparanase-like polypeptide; gene therapy; cancer; angiogenesis;  
 KW trauma; autoimmune disease; skin disease; cardiovascular disease;  
 KW nervous system disease; inflammation; arthritis; genitalia;  
 KW male fertility; erectile dysfunction.

XX Homo sapiens.

FT Key Location/Qualifiers

FT Misc-difference 407 /note= "unspecified residue encoded by KCA"

PN MO200148161-A2.

XX 05-JUL-2001.

XX 18-DEC-2000; 2000MO-EP012309.

XX 23-DEC-1999; 99EP-00125831.

XX (SCHD) SCHERING AG.

XX Siemeister G, Weiss B;

XX WPI; 2001-418259/44.

XX N-PSDB; AAH28347.

PT Human Heparanase-like polynucleotide encoding polypeptides useful for  
 PT modulating expression of the polypeptide and for treating cancer, cancer  
 PT metastasis, aberrant angiogenesis by gene therapy technique.

XX Claim 9; Page 30; 30pp; English.

XX The present sequence represents a human heparanase-like polypeptide.  
 CC Heparanase-like polynucleotides are useful as a source of probes, primers  
 CC and antisense molecules, and in gene therapy. Heparanase-like  
 CC polynucleotides and polypeptides are useful for treating several  
 CC disorders e.g., cancer, cancer metastasis. The oligonucleotides are also  
 CC useful as diagnostic markers for the diagnosis of disorder such as  
 CC cancer, cancer metastasis and aberrant angiogenesis. They may also act as  
 CC diagnostic markers for diagnosis of disorder such as cancer, cancer  
 CC metastasis and aberrant angiogenesis. The heparanase polypeptides and  
 CC polynucleotides are also useful for treating trauma, autoimmune diseases,  
 CC skin diseases, cardiovascular diseases, nervous system diseases, and  
 CC inflammation including arthritis. Since the polynucleotide is  
 CC preferentially expressed in male genitalia, modulation of its expression  
 CC and/or activity may be used for medical intervention in male genitalia  
 CC function that is male fertility control, erectile dysfunction

XX Sequence 492 AA;

Query Match 81.7%; Score 2522; DB 4; Length 492;  
 Best Local Similarity 89.1%; Pred. No. 4e-247;  
 Matches 490; Conservative 0; Mismatches 2; Indels 58; Gaps 1;

QY 43 DRRPLVDRAAGLKEKTLILLDVSTKQPVRTVNFENFSLQDPSIIDDGWLDFSSKRLV 102  
 DB 1 DRRPLVDRAAGLKEKTLILLDVSTKQPVRTVNFENFSLQDPSIIDDGWLDFSSKRLV 60  
 QY 103 TLARGLSPAPLRFGGKRTDPLQFQNLARNPAKSRGGPDYLLKMYEDDIYASDALDKOK 162  
 DB 61 TLARGLSPAPLRFGGKRTDPLQFQNLARNPAKSRGGPDYLLKMYEDDIYASDALDKOK 120  
 QY 163 GCKIAQHPDWMLVLOREKRAQWHVLKKEPSNTYSNLIITARSLLDKLYNFAQSGHLHI 222  
 DB 121 GCKIAQHPDWMLVLOREKRAQWHVLKKEPSNTYSNLIITARSLLDKLYNFAQSGHLHI 161  
 QY 223 PALNALRRNPNNSSSLSLTKYSASKNTNISWELNEBNRYRTMGHRAVNGSOLGKD 282  
 DB 162 -----EPNNRYRTMGHRAVNGSOLGKD 182  
 QY 283 YIQKSLLOPIRITYRSALSYGNIGRPKNVIALIDGFMKVGSTVDVATWQHCVIDGRV 342  
 DB 183 YIQKSLLOPIRITYRSALSYGNIGRPKNVIALIDGFMKVGSTVDVATWQHCVIDGRV 242  
 QY 343 VKVMDFLKTRLDITLSDQIRKIQKVNTTYTPGKKIWEQVYTTAGCTNNISDSYAGFL 402  
 DB 243 VKVMDFLKTRLDITLSDQIRKIQKVNTTYTPGKKIWEQVYTTAGCTNNISDSYAGFL 302  
 QY 403 WNTLGLMANGIDIVIRHSPFDGYNHLVONENPLDPYMLSLIKKELIGPKVLAVHVA 462  
 DB 303 WNTLGLMANGIDIVIRHSPFDGYNHLVONENPLDPYMLSLIKKELIGPKVLAVHVA 362  
 QY 463 GLQKRRPRGVIRDLKLRITAHCTNHHNNYVRSITLFIINLHRSRKKIKLAGTLRDLV 522  
 DB 363 GLQKRRPRGVIRDLKLRITAHCTNHHNNYVRSITLFIINLHRSRKKIKLAGTLRDLV 422  
 QY 523 HOYLLQPYGQGLSKSVQVNGQPLVWDDGTLPELKRPLRAGRTVIRPPTMGFPYVW 582  
 DB 423 HOYLLQPYGQGLSKSVQVNGQPLVWDDGTLPELKRPLRAGRTVIRPPTMGFPYVW 482  
 QY 583 NVNALACRYR 592  
 DB 483 NVNALACRYR 492

## RESULT 12

AA97634 standard; protein; 480 AA.

XX AA97634;

XX 20-APR-2001 (first entry)

DE Human heparanase, hnp1 pns form, protein sequence.

KW Heparanase; hnp1; wound healing; angiogenesis; restenosis; Scrape;  
 KW atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;  
 KW neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;  
 KW gene therapy; human.

XX Homo sapiens.

XX MO200100643-A2.

XX 04-JAN-2001.

XX 19-JUN-2000; 2000MO-IL000358.

XX 25-JUN-1999; 99US-0140801P.

XX (INST-) INSIGHT STRATEGY & MARKETING LTD.

XX Pecker I, Michal I, Itzhaki H;

```

XX  WPI: 2001-137930/14.
DR  N-PSDB; AAA91099.
XX
XX  New polynucleotides and polypeptides that are distantly homologous to
PT  heparanase, useful in wound healing, as well as in gene therapy protocols
PT  for angiogenesis, restenosis, atherosclerosis, or inflammation.
XX
XX  Claim 10; Page 63; 67pp; English.
XX
XX  This sequence represents a heparanase of the invention. The heparanase
CC  DNA and protein sequences are useful in wound healing, angiogenesis,
CC  restenosis, atherosclerosis, inflammation, pulmonary diseases,
CC  neurodegenerative diseases (such as scrapie, Alzheimer's disease, and
CC  Creutzfeldt-Jakob disease) or viral infections. The heparanase coding
CC  sequence is particularly useful in gene therapy
XX
SQ  Sequence 480 AA;
Query Match      79.2%; Score 2446; DB 4; Length 480;
Best Local Similarity 80.9%; Pred. No. 2.1e-239;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;
QY  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDDRAAGLKEKTL 60
DB  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDDRAAGLKEKTL 60
QY  61 ILDDSTKNPVRVVENFLSLQDPSIHDGMDLSKRLVTLARGSPAFIRFGKRT 120
DB  61 ILDDSTKNPVRVVENFLSLQDPSIHDGMDLSKRLVTLARGSPAFIRFGKRT 120
QY  121 DFLQONTNRPNAKSGGPGPDYLLKNYEDDIVRSQVALDKQCKIAQHPDVMVLQREK 180
DB  121 DFLQONTNRPNAKSGGPGPDYLLKNYEDDIVRSQVALDKQCKIAQHPDVMVLQREK 180
QY  121 DFLQONTNRPNAKSGGPGPDYLLKNYEDDIVRSQVALDKQCKIAQHPDVMVLQREK 180
DB  121 DFLQONTNRPNAKSGGPGPDYLLKNYEDDIVRSQVALDKQCKIAQHPDVMVLQREK 180
QY  181 AAQMLVLLKEQFSNTYENLILTLARSIDKLYNPAQSGMLHIFALNMLRRPNNSNNS 240
DB  181 AAQMLVLLKEQFSNTYENLILTLARSIDKLYNPAQSGMLHIFALNMLRRPNNSNNS 240
QY  150 ----- 149
DB  150 ----- 149
QY  241 ALSLIKASAKKYNISWELGNEPNNTRTMGRAVNSQGLQGYVQLKSLQPIRISRAS 300
DB  150 ----- EPNNYRTMGRANVVGSLQGYVQLKSLQPIRISRAS 188
QY  301 LYGPNIGRPKNVIALDGMKVAAGSTVDATVWQHCYDGVVYVMPFLKTRILDTLSDQ 360
DB  189 LYGPNIGRPKNVIALDGMKVAAGSTVDATVWQHCYDGVVYVMPFLKTRILDTLSDQ 248
QY  361 IRKIQKVVNTYTPGKKIMLEGVTTTSACTNNISDSYAAGFLMNTLGMLANQGIIVVIR 420
DB  249 IRKIQKVVNTYTPGKKIMLEGVTTTSACTNNISDSYAAGFLMNTLGMLANQGIIVVIR 308
QY  421 HSFPDHGNHIVDQNFNPLPDYVLSLYKRLIGRPVLAHVAGIQRPRPRVIRDLKRI 480
DB  309 HSFPDHGNHIVDQNFNPLPDYVLSLYKRLIGRPVLAHVAGIQRPRPRVIRDLKRI 368
QY  481 YAHCTNNHNNYVRSITLFTIINLHRSRKKIKLAGTLKDLVHGYLLQPYQGEGLKSKSV 540
DB  369 YAHCTNNHNNYVRSITLFTIINLHRSRKKIKLAGTLKDLVHGYLLQPYQGEGLKSKSV 428
QY  541 QINQGPLVMVDDGTPELKPRPLRAGRTLVPVTMGFFVYKVNALACRIR 592
DB  429 QINQGPLVMVDDGTPELKPRPLRAGRTLVPVTMGFFVYKVNALACRIR 480

```

RESULT 13  
AAU07418  
ID AAU07418 standard; protein; 480 AA.

AAU07418;

18-DEC-2001 (first entry)

Novel human extracellular matrix (ECM) protein #1.

```

XX  Human: immunosuppressive; antiarthritic; antipneumatic; cytosolic;
XX  antiproliferative; cardiac; vasotropic; cerebroprotective; nocotropic;
XX  neuroprotective; antibacterial; virucide; fungicide; ophthalmologic;
XX  extracellular matrix; ECM; autoimmune disease; rheumatoid arthritis;
XX  hyperproliferative disorder; neoplasm; cardiovascular disorder;
XX  cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection;
XX  nervous system disorder; Alzheimer's disease; ocular disorder; sunburn;
XX  wound healing; food additive.
XX
XX  Homo sapiens.
XX
XX  WO200179253-A1.
XX
XX  25-OCT-2001.
XX
XX  11-APR-2001; 2001WO-US011643.
XX
XX  18-APR-2000; 2000US-0198123P.
XX
XX  (HUMA-) HUMAN GENOME SCI INC.
XX
XX  Ftsclla M, Shi Y, Ebner R, Ruben SM;
XX  WPI: 2001-611720/70.
XX  N-PSDB; AAS13843.
XX
XX  New nucleic acids encoding extracellular matrix polypeptides, for
PT  diagnosing, treating, preventing or ameliorating human disorders and
PT  disease, such as, autoimmune, hyperproliferative or cardiovascular
PT  disorders.
XX
XX  Claim 1; Page 292-293; 308pp; English.
XX
XX  The invention relates to novel isolated polynucleotides (I) encoding
CC  extracellular matrix (ECM) polypeptides. (I) and a polypeptide encoded by
CC  (I) are used to prevent, treat or ameliorate a medical condition in e.g.
CC  humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC  are also used in diagnosing a pathological condition or susceptibility to
CC  a pathological condition. The antibodies to the polypeptides can also be
CC  used in alleviating symptoms associated with the disorders and in
CC  diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC  immunoassorbent assays (ELISA). Disorders which are diagnosed or treated
CC  include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC  disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC  e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC  angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC  infections caused by bacteria, viruses and fungi and ocular disorders
CC  e.g. corneal infection. The polypeptides can also be used to aid wound
CC  healing and epithelial cell proliferation, to prevent skin aging due to
CC  sunburn, to maintain organs before transplantation, for supporting cell
CC  culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC  polypeptides can also be used as a food additive or preservative to
CC  increase or decrease storage capabilities. The present sequence
CC  represents the amino acid sequence of novel human extracellular matrix
CC  (ECM) protein #1
XX
SQ  Sequence 480 AA;
Query Match      79.2%; Score 2446; DB 4; Length 480;
Best Local Similarity 80.9%; Pred. No. 2.1e-239;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;
QY  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDDRAAGLKEKTL 60
DB  1 MRVLCAPFPEAMPSSNSRPPACIAPGALYALILHLISSSQAGDRRPLVDDRAAGLKEKTL 60
QY  61 ILDDSTKNPVRVVENFLSLQDPSIHDGMDLSKRLVTLARGSPAFIRFGKRT 120
DB  61 ILDDSTKNPVRVVENFLSLQDPSIHDGMDLSKRLVTLARGSPAFIRFGKRT 120
QY  121 DFLQONTNRPNAKSGGPGPDYLLKNYEDDIVRSQVALDKQCKIAQHPDVMVLQREK 180
DB  121 DFLQONTNRPNAKSGGPGPDYLLKNYEDDIVRSQVALDKQCKIAQHPDVMVLQREK 180

```

```

Db      121 DFLQFQNLRLNPAKSGRGGPPYIYAKNED----- 149
Qy      181 AAQMHVLVLLKEQFSNTYSNLIITARSIDKLYNFADCSGLHIFALNALRRPNNSWNSSS 240
Db      150 ----- 149
Qy      241 ALSLLKYSASKKYNISWELGNEPNRYTMHGRAVNGSLQKDYIQLKSLLOPIRIYSRAS 300
Db      150 -----EPNNYRTMHGRAVNGSLQKDYIQLKSLLOPIRIYSRAS 188
Qy      301 LYGNIGRPRKNVIALLDGFMKVASGYDAVTWQHCYIDGRVYKVMDFLKTLLDITLSDQ 360
Db      189 LYGNIGRPRKNVIALLDGFMKVASGYDAVTWQHCYIDGRVYKVMDFLKTLLDITLSDQ 248
Qy      361 IRKIQRVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLWNTLGMLANOGIDVYIR 420
Db      249 IRKIQRVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLWNTLGMLANOGIDVYIR 308
Qy      421 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHVAAGLQKRPGRVIRDKLRI 480
Db      309 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHVAAGLQKRPGRVIRDKLRI 368
Qy      481 YACCTNHNHNHYVNGSITLFIINLHRSRKKIKLAGTLDKLVHQYLLQPYGEGLSKSV 540
Db      369 YACCTNHNHNHYVNGSITLFIINLHRSRKKIKLAGTLDKLVHQYLLQPYGEGLSKSV 428
Qy      541 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFYVYKVNVALACRYR 592
Db      429 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFYVYKVNVALACRYR 480

RESULT 14
ID      AAB85217 standard; protein; 480 AA.
XX
AC      AAB85217;
DT      07-SEP-2001 (first entry)
XX
DE      Heparanase-like protein Hpa2 splice variant #3.
XX
KM      Heparanase; splice variant; homologue; heparanase-like protein; Hpa2;
KM      cytosolic; neuroprotective; cerebroprotective; immunosuppressive;
KM      antiporter; noctropic; antiinflammatory; antidiabetic; antiaschmatic;
KM      antidiabetic; antiarteriosclerotic; vulnerary.
XX
OS      Homo sapiens.
XX
PN      WO200146392-A2.
XX
PD      28-JUN-2001.
XX
PF      21-DEC-2000; 2000WO-GB004963.
XX
PR      22-DEC-1999; 99GB-00030392.
PR      07-APR-2000; 2000GB-00008713.
XX
PA      (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI      McKenzie EA, Stamps AC, Terrett JA, Tyson KL,
XX
DR      WPI; 2001-418056/44.
XX
DR      N-PSDB; AAH22673.
XX
PT      Novel homologs of heparanase, present in three splice variants, useful
PT      for identifying agents that modulate heparanase, useful in the treatment
PT      and/or prophylaxis of abnormal levels of heparanase.
XX
PS      Claim 1; Fig 3; 97p; English.
XX
CC      The invention provides a homologue to heparanase which is present in
CC      three splice variants. The heparanase homologue polypeptide is useful in
CC      the treatment of a human or non-human animal or for use in diagnosis.

```

```

CC      Vectors comprising the heparanase homologue polynucleotides are useful in
CC      the transformation or transfection of a prokaryotic or eukaryotic host.
CC      The modulators of the polypeptide are useful in the manufacture of a
CC      medicament for the treatment and/or prophylaxis of a condition/disease
CC      associated with abnormal levels of the heparanase homologue, including
CC      cancer, central nervous system (CNS) and neurodegenerative diseases,
CC      cardiovascular diseases such as restenosis following angioplasty and
CC      atherosclerosis, autoimmune diseases, psoriasis, lupus erythematosus,
CC      allografts, inflammatory diseases, arthritis, vascular restenosis, tumour
CC      growth and progression, asthma, Alzheimer's disease, diabetic
CC      retinopathy, wound healing and inflammation. The polypeptide is also
CC      useful in diagnosis and research. The present sequence represents the
CC      amino acid sequence of the smallest splice variant of the heparanase-
CC      like protein Hpa2 of the invention
XX
SQ      Sequence 480 AA;
Query Match      79.2%; Score 2446; DB 4; Length 480;
Best Local Similarity 80.9%; Pred. No. 2,1e-239;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;
Qy      1 MEVLCAFPAMPSNSRPPACIAPGALYIALIHLSSSQAGDRPLPVRAAGLKEKTL 60
Db      1 MEVLCAFPAMPSNSRPPACIAPGALYIALIHLSSSQAGDRPLPVRAAGLKEKTL 60
Qy      61 ILADVSTKNPVATVNEFLSIQLDPSIIHDGMDPLSSKRLVTLARGLSPAFRPGKRT 120
Db      61 ILDDVSTKNPVATVNEFLSIQLDPSIIHDGMDPLSSKRLVTLARGLSPAFRPGKRT 120
Qy      121 DFLQFQNLRLNPAKSGRGGPPDYIYAKNED----- 149
Db      121 DFLQFQNLRLNPAKSGRGGPPDYIYAKNED----- 149
Qy      181 AAQMHVLVLLKEQFSNTYSNLIITARSIDKLYNFADCSGLHIFALNALRRPNNSWNSSS 240
Db      150 ----- 149
Qy      241 ALSLLKYSASKKYNISWELGNEPNRYTMHGRAVNGSLQKDYIQLKSLLOPIRIYSRAS 300
Db      150 -----EPNNYRTMHGRAVNGSLQKDYIQLKSLLOPIRIYSRAS 188
Qy      301 LYGNIGRPRKNVIALLDGFMKVASGYDAVTWQHCYIDGRVYKVMDFLKTLLDITLSDQ 360
Db      189 LYGNIGRPRKNVIALLDGFMKVASGYDAVTWQHCYIDGRVYKVMDFLKTLLDITLSDQ 248
Qy      361 IRKIQRVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLWNTLGMLANOGIDVYIR 420
Db      249 IRKIQRVNTYTPGKKIMLEGVVTTSAAGTNNLSDSYAAGFLWNTLGMLANOGIDVYIR 308
Qy      421 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHVAAGLQKRPGRVIRDKLRI 480
Db      309 HSFEDHGVNHLVDQNFNPLDPYWLSLYKRLIGPKVLAVHVAAGLQKRPGRVIRDKLRI 368
Qy      481 YACCTNHNHNHYVNGSITLFIINLHRSRKKIKLAGTLDKLVHQYLLQPYGEGLSKSV 540
Db      369 YACCTNHNHNHYVNGSITLFIINLHRSRKKIKLAGTLDKLVHQYLLQPYGEGLSKSV 428
Qy      541 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFYVYKVNVALACRYR 592
Db      429 QLNQGPLVWVDGTLPELKPRLRAGRTLVIPVTMGFYVYKVNVALACRYR 480

RESULT 15
ID      AAE18328 standard; protein; 470 AA.
XX
AC      AAE18328;
DT      07-MAY-2002 (first entry)
XX
DE      Human heparanase-2B splice variant protein.
XX
KM      Human; heparanase-2B; Hep-2; wound healing; angiogenesis; restenosis;

```



GenCore Version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:43:17 ; Search time 23 Seconds  
(without alignments)  
1328.808 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088  
Sequence: 1 MRYLCAPPEAMPSSNSRPPA.....PTMGFYVKNVVALACRYR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1165	37.7	536	US-09-930-218-2	Sequence 2, Appl1
2	1163.5	37.7	535	US-09-435-739-44	Sequence 44, Appl1
3	1163.5	37.7	535	US-09-930-218-1	Sequence 1, Appl1
4	1148.5	37.2	543	US-08-922-1708-10	Sequence 10, Appl1
5	1148.5	37.2	543	US-09-071-7398-2	Sequence 2, Appl1
6	1148.5	37.2	543	US-09-260-0388-2	Sequence 2, Appl1
7	1148.5	37.2	543	US-08-635-923-2	Sequence 2, Appl1
8	1148.5	37.2	543	US-09-487-716A-2	Sequence 2, Appl1
9	1148.5	37.2	543	US-09-322-977-2	Sequence 2, Appl1
10	1148.5	37.2	543	US-09-186-200-1	Sequence 1, Appl1
11	1148.5	37.2	543	US-09-435-739-10	Sequence 10, Appl1
12	1148.5	37.2	543	US-09-930-218-3	Sequence 3, Appl1
13	1148.5	37.2	543	US-09-435-739-14	Sequence 14, Appl1
14	1144.5	37.1	588	US-09-181-336-13	Sequence 13, Appl1
15	1144.5	37.1	588	US-09-601-777-2	Sequence 2, Appl1
16	1141.5	37.0	527	US-09-930-218-16	Sequence 16, Appl1
17	1141.5	37.0	530	US-09-252-586-2	Sequence 2, Appl1
18	1104.5	35.8	532	US-09-181-336-15	Sequence 15, Appl1
19	1033.5	33.5	520	US-09-930-218-4	Sequence 4, Appl1
20	984.5	31.9	380	US-09-181-336-19	Sequence 19, Appl1
21	963.5	31.2	380	US-09-181-336-17	Sequence 17, Appl1
22	1144	4.7	38	US-09-252-586-26	Sequence 26, Appl1
23	114.5	3.7	399	US-08-282-197C-54	Sequence 54, Appl1
24	114.5	3.7	837	US-09-390-234-14	Sequence 14, Appl1
25	114.5	3.3	1183	US-09-603-311-14	Sequence 14, Appl1
26	100.5	3.3	1183	US-09-532-3108-6	Sequence 6, Appl1
27	97.5	3.2	774	US-09-328-352-5361	Sequence 5361, Ap

28	96.5	3.1	1874	US-09-331-403-2	Sequence 2, Appl1
29	96	3.1	990	US-08-392-625-20	Sequence 20, Appl1
30	96	3.1	990	US-08-466-961A-20	Sequence 20, Appl1
31	95.5	3.1	709	US-09-668-673B-3	Sequence 3, Appl1
32	95.5	3.1	849	US-08-405-392-18	Sequence 18, Appl1
33	95.5	3.1	849	US-08-487-691-18	Sequence 18, Appl1
34	95.5	3.1	849	US-08-189-738A-19	Sequence 19, Appl1
35	95.5	3.1	866	US-08-405-392-17	Sequence 17, Appl1
36	95.5	3.1	866	US-08-487-691-17	Sequence 17, Appl1
37	95.5	3.1	893	US-08-189-738A-18	Sequence 18, Appl1
38	95.5	3.1	905	US-08-405-392-2	Sequence 2, Appl1
39	95.5	3.1	905	US-08-487-691-2	Sequence 2, Appl1
40	95.5	3.1	905	US-08-666-221B-4	Sequence 4, Appl1
41	95.5	3.1	905	US-08-666-221B-10	Sequence 10, Appl1
42	95.5	3.1	905	US-08-189-738A-2	Sequence 2, Appl1
43	95	3.1	371	US-09-758-008-1	Sequence 1, Appl1
44	94	3.0	1103	US-08-455-543A-53	Sequence 53, Appl1
45	94	3.0	1103	US-08-223-305C-53	Sequence 53, Appl1

#### ALIGNMENTS

```

RESULT 1
US-09-930-218-2
Sequence 2, Application US/09930218
Patent No. 6677137
GENERAL INFORMATION:
APPLICANT: Goldschmidt, orit
APPLICANT: pecker, iris
APPLICANT: vlodavsky, israel
APPLICANT: israel, michael
TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H
FILE REFERENCE: 01/22335
CURRENT APPLICATION NUMBER: US/09/930, 218
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/666, 390
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 536
TYPE: PRT
ORGANISM: Rattus rattus
US-09-930-218-2

Query Match      37.7%; Score 1165; DB 4; Length 536;
Best Local Similarity 45.1%; Pred. No. 1e-11;
Matches 248; Conservative 87; Mismatches 187; Indels 28; Gaps 7;

QY 42 GRRRLPVDRAAG-LKEKTLILDVSTKPNVRYNENFLSLQDPSIHD-GMIDFLSSK 99
DB 12 GRLRLTGTPGNAPTVDVDELYTKRLFOQSVSPSLITIASLATDRFLFLGSP 71
QY 100 RLVTLARGLSPAFLFGSKRTPLOFONLRNPAYSRGGPGDYLYKAYEDDIVASVALD 159
DB 72 RLALARGLSPAYLRFQGTXTDFLIF-----DPNKEPTSEERSYVQSDNNICSE----- 123
QY 160 KQKGGKIQHPDVMVLQREKAQMHTVLKEQNSNTSNLILARSLDKYNFADCSGL 219
DB 124 -----RVSAVLEKQMEWPFQ-ELHLLEQYQREPKNSTYRSYDMLYSFAKSRLL 175
QY 220 HLIFALNLRNPNNSWSSALSLTKYSASKYNIISWELNENPNYRTMGRAVNSQL 279
DB 176 DIFELNLLRPPDRMSSNAQLLINCSKYNISWELNENPNSEFWKKAQISIDGL 235
QY 280 GGDVYQLSLQPIRTYERASLYGPNIGRPRKNYIALDGMKYVAGSYVDVATMOHCYTD 339
DB 236 GDFVELHLKLOK-SAFONAKLYGPDIGQPRKTVKLRSFLKAGGEVYDLSLTHHLYLN 294
QY 340 GGVVVMDFLKLRLDITLSDQIRKIOKVNTYTPGKTIWLEGVVYTSAGTNNLSDSYAA 399

```

```

Db 295 GRVATXEDFLSSDVLDTFLILSVOKILKVTKEMTGPKKWLGETSSAYGGAPLSTFPA 354
Qy 400 GFLMLNTLGLANOGIDVVRHSFDFHGVNHLVDONPNP:PDYMLSLYKRLIGPVAV 459
Db 355 GFWMMDLGLSAQGLLEVVRQVFFGAGNHLVDENFEP:PDYMLSLYKRLIGPVAV 414
Qy 460 HVAGLQKRPGRVIRDKRIYAHCTNNHNNYVRSITLFIINLHRSRKKIKLAGTLD 519
Db 415 RVKGPD-----RSKLRVYLHCTVNVHPRYREGDLTYVNLNHNVTGHLKLPFPMFS 465
Qy 520 KLHVHVLQPYGOGKSKSVOLNGOPLVWDDGT:PELKRPRLRAGRTLVIPVPMGFY 579
Db 466 RPYDKILKFPSSDGLSKSVOLNGOPLVWDDGT:PELKRPRLRAGRTLVIPVPMGFY 525
Qy 580 VVKNVNALAC 589
Db 526 VIRNAKIAAC 535

```

## RESULT 2

```

US-09-435-739-44
; Sequence 44, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Pecker, Itis
; APPLICANT: Vlodavsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435, 739
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent version 3.0
; SEQ ID NO 44
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-435-739-44

```

```

Query Match 37.7%; Score 1163.5; DB 4; Length 535;
Best Local Similarity 46.2%; Pred. No. 1.4e-111;
Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

Qy 63 LDVSTKNFRTVNNFSLQDPSIHD-GWLDPLSSKRLVTLARGLSPAFIRFGKRTD 121
Db 33 LEFTKRLRSLVSPFSLITIDAGLATDPRFLTGLSPRLALRGLSPAYLRFGGTKTD 92
Qy 122 FLQPNLRNPAKSGGPGPDYILKNYEDDIYRSVVALDKQKCKIAQHPDVMVLQREKA 181
Db 93 FLIF---DPKEPTSEERSYKWSQVNHDCRSEPV-----SAALRLQVEMP 137
Qy 182 AQMLVLKQEPSTYNSLILTARSLDKLYNFADCSGLHLIFALNALRRPNNSWSSSA 241
Db 138 FQ-ELLRLRQYQKEFNSTYSRSVDMLYSFACSGDLDFGLNALLRTPDLRMSNA 196
Qy 242 LSLKYSASKKYNISWELGNEPNRYRTMGRVANGSOLQGYIQKSLQPIIYRSASL 301
Db 197 QLLDLYCSSKYNISWELGNEPNRPMKKAHLIDGLQGBDFVFLHKLQK-SAFQNAKL 255
Qy 302 YGNIGRPRKRVIALLDGFMKVASGTVDVAVTMOHCYIDGRVVKVMDFLKRLDLSDOI 361
Db 256 YGPRIGQKRGKTVLRLSFLKAGGEVIDSLTMHHYVINGRIATKEDFLSSDALDTFLSV 315
Qy 362 RKIQVNVNTYTPGKKIWELEGVVTTSSAGTNNLSDSYAGFLMNTLGLANOGIDVYIRH 421
Db 316 QKILKVTKEITPPGKKVWLGERTSSAYGGAPLSTFPAAGFMWLDKLGISAQMGIEVVRQ 375
Qy 422 SFPDGYNHLVDONFNP:PDYMLSLYKRLIGPVAVLHVAGLQKRPGRVIRDKLRIY 481
Db 376 VFFGAGNHLVDENFEP:PDYMLSLYKRLIGPVAVLSRVKGPD-----RSKLRVY 426
Qy 482 AHCTNNHNNYVRSITLFIINLHRSRKKIKLAGTLDKLVHGYLLQPYGQGLSKSVQ 541

```

```

Db 427 LHCTVNVHPRYQGBDRLTYVNLNHNVTGHLKVPPELPRKKVDYTLKPSGPDGLSKSVQ 486
Qy 542 LNGOPLVWDDGT:PELKRPRLRAGRTLVIPVPMGFYVKNVNALAC 589
Db 487 LNGOPLVWDDGT:PELKRPRLRAGRTLVIPVPMGFYVKNVNALAC 534

```

## RESULT 3

```

US-09-930-218-1
; Sequence 1, Application US/09930218
; Patent No. 6677137
; GENERAL INFORMATION:
; APPLICANT: goldsmith, orit
; APPLICANT: pecker, itis
; APPLICANT: vlodavsky, israel
; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 01/22335
; CURRENT APPLICATION NUMBER: US/09/930, 218
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/666, 390
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-930-218-1

```

```

Query Match 37.7%; Score 1163.5; DB 4; Length 535;
Best Local Similarity 46.2%; Pred. No. 1.4e-111;
Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

Qy 63 LDVSTKNFRTVNNFSLQDPSIHD-GWLDPLSSKRLVTLARGLSPAFIRFGKRTD 121
Db 33 LEFTKRLRSLVSPFSLITIDAGLATDPRFLTGLSPRLALRGLSPAYLRFGGTKTD 92
Qy 122 FLQPNLRNPAKSGGPGPDYILKNYEDDIYRSVVALDKQKCKIAQHPDVMVLQREKA 181
Db 93 FLIF---DPKEPTSEERSYKWSQVNHDCRSEPV-----SAALRLQVEMP 137
Qy 182 AQMLVLKQEPSTYNSLILTARSLDKLYNFADCSGLHLIFALNALRRPNNSWSSSA 241
Db 138 FQ-ELLRLRQYQKEFNSTYSRSVDMLYSFACSGDLDFGLNALLRTPDLRMSNA 196
Qy 242 LSLKYSASKKYNISWELGNEPNRYRTMGRVANGSOLQGYIQKSLQPIIYRSASL 301
Db 197 QLLDLYCSSKYNISWELGNEPNRPMKKAHLIDGLQGBDFVFLHKLQK-SAFQNAKL 255
Qy 302 YGNIGRPRKRVIALLDGFMKVASGTVDVAVTMOHCYIDGRVVKVMDFLKRLDLSDOI 361
Db 256 YGPRIGQKRGKTVLRLSFLKAGGEVIDSLTMHHYVINGRIATKEDFLSSDALDTFLSV 315
Qy 362 RKIQVNVNTYTPGKKIWELEGVVTTSSAGTNNLSDSYAGFLMNTLGLANOGIDVYIRH 421
Db 316 QKILKVTKEITPPGKKVWLGERTSSAYGGAPLSTFPAAGFMWLDKLGISAQMGIEVVRQ 375
Qy 422 SFPDGYNHLVDONFNP:PDYMLSLYKRLIGPVAVLHVAGLQKRPGRVIRDKLRIY 481
Db 376 VFFGAGNHLVDENFEP:PDYMLSLYKRLIGPVAVLSRVKGPD-----RSKLRVY 426
Qy 482 AHCTNNHNNYVRSITLFIINLHRSRKKIKLAGTLDKLVHGYLLQPYGQGLSKSVQ 541
Db 487 LNGOPLVWDDGT:PELKRPRLRAGRTLVIPVPMGFYVKNVNALAC 589

```

RESULT 4  
US-08-922-170B-10  
Sequence 10, Application US/08922170B  
Patent No. 5968822  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker, Israel Vlodavsky and Elena  
Applicant: Feinstein  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
TITLE OF INVENTION: HAVING HEPARANASE ACTIVITY AND EXPRESSION OF  
TITLE OF INVENTION: SAME IN TRANSFUSED CELLS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Robert Sheindein  
STREET: 2940 Birchtree lane  
CITY: Silver Spring  
STATE: Maryland  
COUNTRY: United States of America  
ZIP: 20906  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimnote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
SOFTWARE: an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/922,170B  
FILING DATE: 2 SEP 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-5625553  
TELEFAX: 972-3-5625554  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-922-170B-10

Query Match 37.2%; Score 1148.5; DB 2; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110; Indels 53; Gaps 9;  
Matches 249; Conservative 82; Mismatches 190;  
Query 18 PPACIAPGALYALALLHLISLSQAGDRRLPVDRRAAGLKEKTLILLDVSTKQNPRTVVEN 77  
Db 20 PLGLPSGAL-----PRPA-----QAQDVVDLDFEQEHLHVSFS 55  
Query 78 FLISQOLDPSIHD-GWIDFLSSKRLVTLARGSLPAFLREGKXTDFQFQNLNPAKSRG 136  
Db 56 FLSTVTDANLADPRLFLILGSPPLRLTARGLSAVYRFGYKIDFLIF---DPKKEST 111  
Query 137 GPGDYYLYKNYBDDIVSDVALDKQKCKIAQ-HPDVMTVLOREKAQOMLVLLKQPSN 195  
Db 112 FEERSYVQSQVNDI-----CKYGSIPDVEEKLRLIEWPYQEGT-LLEHYOK 158  
Query 196 TYSVLIITARSIDLCYNFADCSGLHIFALNALRNNNNSSSSALSILKYASKKYN 255  
Db 159 KFKKSTYSRSSVDVLYTFANCSSGLDIFGNALRLRTADLQWNSNNAQLLDYCSSKGYNI 218  
Query 256 SWEIENPNRYRTMGRAVNGSOLGKDYIOLKSLIOPRIRYSRSLVGNIGRPKRVIA 315  
Db 219 SWEIENPNRSLKADLFINGSQLGEDIYIOHKLRLK-STFKAKAKLYGPGVCGPRRTAK 277  
Query 316 LLDSFMKVAGSTVDATVWQHGYIDGRVYKWMDFLKTRLDLTSLDQIKKIQVNTYTPGK 375

Db 278 MLKSFPLKAGGEVIDSVTHHHYILNGRATREDPLNPVDLIFLSSVQKQVVESTRPGK 337  
Query 376 KIWLEGVATTSGAGTNNLSQSYAAGF.MLNTLGMLANQIGIDVIRASFPHGNTNLYDN 435  
Db 338 KVMLEGTSSAVGGAPLISDTFAAGFMWLDRLG.SAMGIEVYWRQVFFAGVYHLVDEN 397  
Query 436 ENPLPDYMLSLYRELLGPKYLAHVAGLQKPRGPRVIRDKLRIYACTNHHNHYVVG 495  
Db 398 FDLPLDYMSLFLFKVLGTYKVLMAVSGSKRR-----KRVYHACTNTDNPRIKGG 448  
Query 496 SITLEFINLHRSRRKIKIAGTLRDKLVHGYILOPYGEGKSKSVQNLQPLVYVDDGTL 555  
Db 449 DLTLYAINLHVNTKYLRLPYFPNSKQVDKYLRLPLGPHGLSKSVQNLGLTLKXVDDTL 508  
Query 556 PELKPRPLRAGRTLVIPYTMGPFYVNVNVALAC 589  
Db 509 PPLMEKPLRPGSSGLPLPFSYSPFVIRNAKYAAC 542

RESULT 5  
US-09-071-739B-2  
Sequence 2, Application US/09071739B  
Patent No. 617545  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker et al.  
TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL  
TITLE OF INVENTION: APPLICATIONS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Anthony Caetorina  
STREET: 20001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimnote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,739B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-5625553  
TELEFAX: 972-3-5625554  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-071-739B-2  
Query Match 37.2%; Score 1148.5; DB 3; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;  
Query 18 PPACIAPGALYALALLHLISLSQAGDRRLPVDRRAAGLKEKTLILLDVSTKQNPRTVVEN 77



```

Db      20 FLGPISEPAL-----PRPA-----QADVVLDLDFPTOEPLHVS 55
Qy      78 FLSLODPSIIHD-GWLDPLSKRLVTLARGLSPAPLRFSGKRTDLOFONLRNPASRG 136
Db      56 FLSTVIDANLMTDPRFILLGSPKRLRLARGLSPAYLRFSGTKDPLIF----DPKEEST 111
Qy      137 GPGPDYLYKATVEDDLYRSVDALDKQCKCKIAQ-HPDYMVLVQREKAAQMHLYLKQFSN 195
Db      112 FEERSYQSQVNQDI-----CKYGSIPDVEEKLRLWMPYQEQ-LRREHYOK 158
Qy      196 TYSNLIITARSIDLKLYNFADCSGLHLIFALNALRPNPNNSWSSSALSLKYASAKKYN 255
Db      159 KFKNSTYRSRSVDLYTFANCSGDLIFGLNALRTADLQWNSNAGLLDYCSSKGYNI 218
Qy      256 SWEIGNEPNRYTMHGRAVNSQGLKDYIOLKSLLOPRIRYSPASLYGPNIGRPRKNVIA 315
Db      219 SWEIGNEPNRSFLKKADIFINGSQIGEDYIOLHKLRLK-STFKNAKLYGPDVGQPRKRTAK 277
Qy      316 LLDGFEMKAVGSTVDATVMOHCYIDGRVVKWMDFLKTRLLDPLSDQIRKIQKVNNTYTPGK 375
Db      278 MLKSFPLKAGGEVIDSVTHMYTLNKRATREDPLNPVDLIFISSVQKVFQVESTREPGK 337
Qy      376 KIMLEGVTTTSAGGTNNISDSYAGFLMNTLQMLANGIDVYIRHSFEDHGYNHLVDON 435
Db      338 KYWLGETSSAYGGAPLLSDTFPAAGFWMLDKLGASAMGIEVVMQVFFGAGNYHLVDEN 397
Qy      436 FNLPLDPMILLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRIYACNHNHNHYRG 495
Db      398 FDLPLDPMILSLFKLVGKYLKVAASVQSKR-----KLRVYLCTNDNPRYKEG 448
Qy      496 SITFLIINLHRSRKIKLAGTRDKLVHOYLLQPYGQGLSKSVQNLGOLPLWMDGTL 555
Db      449 DLTLYAINLHNTYKXRLRFPFNSKQYDKYLRLRPHGLSKSVQNLGTLKWDQTL 508
Qy      556 PELKPRPLRAGRTIVIPVTMGFYVKNVNALAC 589
Db      509 PLMEKPLRPGSSLGLPARSYSFVIRNAKVAAC 542

RESULT 6
US-09-260-038B-2
; Sequence 2, Application US/09260038B
; Patent No. 6348344
; GENERAL INFORMATION:
; APPLICANT: Maly Ayal-Hershkovitz et al.
; TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR
; EXPRESSING RECOMBINANT HEPARANASE
; AND METHODS OF PURIFYING SAME
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead+ Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted to
; an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/260,038B
; FILING DATE: 02-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/071,618
; FILING DATE: May 1, 1998
; APPLICATION NUMBER: 09/071,739
; FILING DATE: May 1, 1998

```

```

; APPLICATION NUMBER: 08/922,180
; FILING DATE: September 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 910/16
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-562553
; TELEFAX: 972-3-562554
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-260-038B-2

Query Match      37.2%; Score 1148.5; DB 4; Length 543;
Best Local Similarity 43.4%; Pred. No. 5.3e-110;
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

Qy      18 PPACIAPGALVIALLIHLSSQAGDRRPLVDRAAGLKEKTLILDVSTKNIPVRYTEN 77
Db      20 FLGPISEPAL-----PRPA-----QADVVLDLDFPTOEPLHVS 55
Qy      78 FLSLODPSIIHD-GWLDPLSKRLVTLARGLSPAPLRFSGKRTDLOFONLRNPASRG 136
Db      56 FLSTVIDANLMTDPRFILLGSPKRLRLARGLSPAYLRFSGTKDPLIF----DPKEEST 111
Qy      137 GPGPDYLYKATVEDDLYRSVDALDKQCKCKIAQ-HPDYMVLVQREKAAQMHLYLKQFSN 195
Db      112 FEERSYQSQVNQDI-----CKYGSIPDVEEKLRLWMPYQEQ-LRREHYOK 158
Qy      196 TYSNLIITARSIDLKLYNFADCSGLHLIFALNALRPNPNNSWSSSALSLKYASAKKYN 255
Db      159 KFKNSTYRSRSVDLYTFANCSGDLIFGLNALRTADLQWNSNAGLLDYCSSKGYNI 218
Qy      256 SWEIGNEPNRYTMHGRAVNSQGLKDYIOLKSLLOPRIRYSPASLYGPNIGRPRKNVIA 315
Db      219 SWEIGNEPNRSFLKKADIFINGSQIGEDYIOLHKLRLK-STFKNAKLYGPDVGQPRKRTAK 277
Qy      316 LLDGFEMKAVGSTVDATVMOHCYIDGRVVKWMDFLKTRLLDPLSDQIRKIQKVNNTYTPGK 375
Db      278 MLKSFPLKAGGEVIDSVTHMYTLNKRATREDPLNPVDLIFISSVQKVFQVESTREPGK 337
Qy      376 KIMLEGVTTTSAGGTNNISDSYAGFLMNTLQMLANGIDVYIRHSFEDHGYNHLVDON 435
Db      338 KYWLGETSSAYGGAPLLSDTFPAAGFWMLDKLGASAMGIEVVMQVFFGAGNYHLVDEN 397
Qy      436 FNLPLDPMILLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRIYACNHNHNHYRG 495
Db      398 FDLPLDPMILSLFKLVGKYLKVAASVQSKR-----KLRVYLCTNDNPRYKEG 448
Qy      496 SITFLIINLHRSRKIKLAGTRDKLVHOYLLQPYGQGLSKSVQNLGOLPLWMDGTL 555
Db      449 DLTLYAINLHNTYKXRLRFPFNSKQYDKYLRLRPHGLSKSVQNLGTLKWDQTL 508
Qy      556 PELKPRPLRAGRTIVIPVTMGFYVKNVNALAC 589
Db      509 PLMEKPLRPGSSLGLPARSYSFVIRNAKVAAC 542

RESULT 7
US-09-635-923-2
; Sequence 2, Application US/09635923
; Patent No. 6426209
; GENERAL INFORMATION:
; APPLICANT: Maly Ayal-Hershkovitz et al.
; TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR
; EXPRESSING RECOMBINANT HEPARANASE
; AND METHODS OF PURIFYING SAME

```

NUMBER OF SEQUENCES: 25.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: PC  
 OPERATING SYSTEM: MS DOS version 6.2,  
 Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/635,923  
 FILING DATE: 10-Aug-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/487,716  
 FILING DATE: 19-Jan-2000  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-5625553  
 TELEFAX: 972-3-5625554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 543  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-635-923-2

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
 Best Local Similarity 43.4%; Pred. No. 5.3e-110; Indels 53; Gaps 9;  
 Matches 249; Conservative 82; Mismatches 190;

18 PPAICAPALYALILHLISLSQAQDRRLPVDRRAAGLKEKTLILLDVTSTNPRVTNEN 77  
 20 PLGLPSGAL-----PRPA-----QAQDVVDLDFEFTQEPHLVSPS 55  
 78 FLISOLDPSIIHD-CWLDPLSKRLVTLARGSPAFLEFGGRKDPLOFQMLRPAKSRG 136  
 56 FLSTVIDANLTDPFLLILGSPKRTIARGLSPAYLRFGGTKIDFLF---DPKKEST 111  
 137 GGPDPYLYKNYEDDIVRSVALDKQCKKIAQ-HEDVNLVIQREKPAQMLVLLKEQFSN 195  
 112 FEERSYQSQVNOPI-----CKYGSIPDVBEKRLLEPYEQCL-LREHYOK 158  
 196 TYSNLIILARSLDLYNADCSGLHLIPALNALRPNPNNSNSALSILKYSKCKNI 255  
 159 KFKNSTYRSRSDVLYTFANCSGDLIFGLNALFTALQNMSSNAQCLLDYCSSKGINI 218  
 256 SWELENEPNRYTRHGRAVNGSOLGKDYIOLKSLIQIRIYSRSLYQGNIGRRKKNYA 315  
 219 SWELENEPNRSLFKADIFINGSGEDYIQLHKLIRK-STFENAKLYGPDVGQRRKRTAK 277  
 316 LIDGMAKTAGSTVDVATWQCHTIDRVRVAKMDFKTRILDLTSDIRIKQVNTYTPGK 375  
 278 MLKSTLKAGGEYIDVTVTHHYLNGRTAIFREDFLNPVDLDFISSVQKVFQVVESTTRPK 337  
 376 KIMLEGVVTSAGTNNLSDSYAAQFLWMLNTLGLMANOGIDVIVIRHSFPHGHVNLVDON 435

338 KWLMEETSSAYGGAPLLSDTFPAAGFWMLDKLGSAENGIVNWRQVFFGAGNVLVDEN 397  
 436 FNPPLDYMILSLYKALIGPKYLAHVAGLQRRKRGVIRPKLRIYACTHHNNHYRG 495  
 398 FDPPLDYMILSLFKLVGTGKILMASVQSKR-----KLVYLLHCTGTNDPRYMEG 448  
 496 STTLFTIINLRERKKIKIAGTRDKLVHGYLLQPYGEGLSKSVQNLGQPLWVYDQTL 555  
 449 DLTLYAINHNVTKYRLPYFPNKKQVDRKYLRLPGFPHGLLSKSVQNLGTLKWDQTL 508  
 556 PELKRPRLAGRTVLPVWGMFYVKNVALAC 589  
 509 PLMEKPLRPSGLGDPASISFFVIRAKYAAAC 542

RESULT 8  
 US-09-487-716A-2  
 Sequence 2, Application US/09487716A  
 Patent No. 6475763  
 GENERAL INFORMATION:  
 APPLICANT: Macy Ayal-Herskovitz et al.  
 TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
 EXPRESSING RECOMBINANT HEPARINASE  
 AND METHODS OF PURIFYING SAME  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: PC  
 OPERATING SYSTEM: MS DOS version 6.2,  
 Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/487,716A  
 FILING DATE: 19-Jan-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-5625553  
 TELEFAX: 972-3-5625554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 543  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-487-716A-2

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
 Best Local Similarity 43.4%; Pred. No. 5.3e-110; Indels 53; Gaps 9;  
 Matches 249; Conservative 82; Mismatches 190;

18 PPAICAPALYALILHLISLSQAQDRRLPVDRRAAGLKEKTLILLDVTSTNPRVTNEN 77

Fri May 14 13:01:55 2004

us-10-088-676-2.rat

Page 6

Db 20 PLGLSPGAL-----PRPA-----QAQDVVDLDFTQEPHLVSPS 55  
Qy 78 FLSTLDPSIIHD-GMLDFLSKRLVTLARGLSPAFIRFGKRTDPLQFONLRNPAKSRG 136  
Db 56 FLSTVIDANLATDPRFILLGLSPKRLTARGLSPAYIRFGSTKTDFLIF----DPKKEST 111  
Qy 137 GPSPDYLLKNEEDIVRSVALDKQKCKIAQ-HPDVMVLQREKAAQMHVLVLLKEQFSN 195  
Db 112 FEERSYMOQVNOI-----CKYGSIPPDVEKRLRLEMPYQEOQL-LIREHYOK 158  
Qy 196 TYSNLILPARKSLDLYNFADCSGLHLIFALNALRPNPNSSSALSLKYSASKKYN 255  
Db 159 KFNKSTYSRSSVDLYTFPANCSSGLDLIFGLNALRTADLQNSSNAQLLDYCSSKGYNI 218  
Qy 256 SWEIGNEPNRYRTMGRAVNGSOLGKDYIQKSLQPIRISRAISLYGNIGPRKKNVIA 315  
Db 219 SWEIGNEPNRSFLKADIFINSGQLGEDIYQHLKLRK-STFKNAKLYGPDVGQPRKRTAK 277  
Qy 316 LLDGFMKVAGSTVDAYTMQHCYIDGRVYKVMDFLKTLLDPLSDQIRKIQVNTYTPGK 375  
Db 278 MLKSFELKAGGEVIDSVTMHRYTLNGRTATREDFLNPVDLIFISVQKVFQVSESTRPGK 337  
Qy 376 KIMLEGVTTTSGGTNNLSDSYAAGFLMNTLGMLANQIGIDVYIRHSFPDHGYNHLVDON 435  
Db 338 KVMLGSTSSAYGAGAPLSDTFAAGFWMLDKLGASBMGIEVMVRQVFFGAGNYHLVDEN 397  
Qy 436 FNPDPYWLSLLYKRLIGPKVLAVHAGLQKPRPGVIRDKLRIYACHTNNHNNVYRG 495  
Db 398 FDPDPYWLSLFLKLVGTVKVLMAVQSSKRR-----KLRVYHCTNTDPRYKRG 448  
Qy 496 SITFLIINHRSRKIKLAGTLRDLYVQYLLQPYGQGLSKSVQUNGOPLVWVDGTL 555  
Db 449 DLTVALINLHNTYKTLRPLPFPSSNKQVDKYLRLPGHGLSKSVQUNGTLTKWVDQTL 508  
Qy 556 PELKPRPLRAGRTVLPVYTMGFYVVKVNAALAC 589  
Db 509 PLMEKPLRPGSSLGLPFSYSFFVIRNAKVAAC 542

RESULT 9  
US-09-322-977-2  
Sequence 2, Application US/09322977  
Patent No. 6531129

GENERAL INFORMATION:

APPLICANT: Iris Pecker et al.

TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES

TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL

TITLE OF INVENTION: APPLICATIONS

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 2001 Jefferson Davis Highway, Suite 207

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead\* Slimote-890TX

OPERATING SYSTEM: MS DOS version 6.2,

SOFTWARE: Word for Windows version 3.11

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/322,977

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/922,180

FILING DATE: September 2, 1997

APPLICATION NUMBER: 09/071,739

FILING DATE: May 1, 1998

ATTORNEY/AGENT INFORMATION:

NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-322-977-2

Query Match 37.2% Score 1148.5; DB 4; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

Db 18 PPACIAPGALYALALHLHLSSQAQDRRPLFVDBAAGLKEKTLILLDVSTKNPVRTNEN 77  
Qy 20 PLGLSPGAL-----PRPA-----QAQDVVDLDFTQEPHLVSPS 55  
Db 78 FLSTLDPSIIHD-GMLDFLSKRLVTLARGLSPAFIRFGKRTDPLQFONLRNPAKSRG 136  
Db 56 FLSTVIDANLATDPRFILLGLSPKRLTARGLSPAYIRFGSTKTDFLIF----DPKKEST 111  
Qy 137 GPSPDYLLKNEEDIVRSVALDKQKCKIAQ-HPDVMVLQREKAAQMHVLVLLKEQFSN 195  
Db 112 FEERSYMOQVNOI-----CKYGSIPPDVEKRLRLEMPYQEOQL-LIREHYOK 158  
Qy 196 TYSNLILPARKSLDLYNFADCSGLHLIFALNALRPNPNSSSALSLKYSASKKYN 255  
Db 159 KFNKSTYSRSSVDLYTFPANCSSGLDLIFGLNALRTADLQNSSNAQLLDYCSSKGYNI 218  
Qy 256 SWEIGNEPNRYRTMGRAVNGSOLGKDYIQKSLQPIRISRAISLYGNIGPRKKNVIA 315  
Db 219 SWEIGNEPNRSFLKADIFINSGQLGEDIYQHLKLRK-STFKNAKLYGPDVGQPRKRTAK 277  
Qy 316 LLDGFMKVAGSTVDAYTMQHCYIDGRVYKVMDFLKTLLDPLSDQIRKIQVNTYTPGK 375  
Db 278 MLKSFELKAGGEVIDSVTMHRYTLNGRTATREDFLNPVDLIFISVQKVFQVSESTRPGK 337  
Qy 376 KIMLEGVTTTSGGTNNLSDSYAAGFLMNTLGMLANQIGIDVYIRHSFPDHGYNHLVDON 435  
Db 338 KVMLGSTSSAYGAGAPLSDTFAAGFWMLDKLGASBMGIEVMVRQVFFGAGNYHLVDEN 397  
Qy 436 FNPDPYWLSLLYKRLIGPKVLAVHAGLQKPRPGVIRDKLRIYACHTNNHNNVYRG 495  
Db 398 FDPDPYWLSLFLKLVGTVKVLMAVQSSKRR-----KLRVYHCTNTDPRYKRG 448  
Qy 496 SITFLIINHRSRKIKLAGTLRDLYVQYLLQPYGQGLSKSVQUNGOPLVWVDGTL 555  
Db 449 DLTVALINLHNTYKTLRPLPFPSSNKQVDKYLRLPGHGLSKSVQUNGTLTKWVDQTL 508  
Qy 556 PELKPRPLRAGRTVLPVYTMGFYVVKVNAALAC 589  
Db 509 PLMEKPLRPGSSLGLPFSYSFFVIRNAKVAAC 542

RESULT 10  
US-09-186-200-1  
Sequence 1, Application US/09186200  
Patent No. 6562950

GENERAL INFORMATION:

APPLICANT: Peretz, Tuva et al.

TITLE OF INVENTION: HEPARANASE ACTIVITY NEUTRALIZING ANTI-HEPARANASE MONOCLONAL ANTIBIC

TITLE OF INVENTION: APPLICATIONS

NUMBER OF SEQ ID NOS: 1

SOFTWARE: Patentin version 3.1

SEQ ID NO 1

LENGTH: 543

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-186-200-1

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

18 PPACIAGALYIALLLHLSSQAGDRRLPVDRRAAGLKEKTLILDVSTKRVRYTEN 77  
DB PLGLSPGAL-----PRPA-----QAQDVVDLDFFTQEPHLVSPS 55  
QY FLSLQDPSIITH-GWLDPLSSKRLVTLARGLSPAFIRFGGKRTDFLOFQNLNPAKSRG 136  
DB FLSTVIDANLATDPRFLILGSPKLRRLARGLSPAYIRFGGKTDFLIF---DPKEST 111  
QY 137 GPGPDYLYKNVEDIVRSQVADLQKQCKIAQ-HPDVMVLQREKAQOMELVLLKEQFSN 195  
DB 112 FEERSYMQSQVNODI-----CKYGSIPPDVEEKLRLMPYQEQ-LIREHYQK 158  
QY 196 TYSNLIILFASLIDKLYNFADCGSLHIFALNALRRNNNNSSSSALSLKYSASKKYN 255  
DB 159 KFKNSTYRSRSDVLYITFANCSGDLIFGLNALRTADLQWSSNAQLLDYCCKSGYNI 218  
QY 256 SWEIGNEPNRYTMMGRAVNGSQLGKDYIQLKSLQPIRTYRSALYGPNIQPRKXVIA 315  
DB 219 SWEIGNEPNRYTMMGRAVNGSQLGKDYIQLKSLQPIRTYRSALYGPNIQPRKXVIA 315  
QY 316 LLDGFMKVAAGSTVDAVTWQHCVYIDGRVYKVMDFLKTLLDLSQIRKIQKVNTYTPGK 375  
DB 278 MLKSTFKAGGEVIDSVTHHYLYNKRTRATREDFLNPVDLIFISSQKVPQVSESTRPK 337  
QY 376 KIMLEGVVTTSAGTNNISDSYAAGFLMNTLGMLANOGIDVIRHSFFDHGYNHLYDON 435  
DB 338 KWLGETSSAYGGAPLSDITFAAGFMWLDKLGSAEMGIEVMQVFFGAGNTHLVDEN 397  
QY 436 ENPLPDYMLSLYKLLIGPKYLAHVAGLQKRPGRVIRDKLRIYACNTHNNHNYVVG 495  
DB 398 FDPDPYMLSLFVKLVGTQVLMASVQSKRR-----KLRYVLCNTNDNRYKEG 448  
QY 496 SITFLIINLHRSRKKIKLAGTRDLYVHGYLLQPYGQELSKSVQVINGQPLVWDDGT 555  
DB 449 DLTVAIINLHNTKYLRLPYFPNSKQVQDKYLLRPLGPHGLSKSVQVINGQPLVWDDGT 508  
QY 556 PELKRPRLRAGRTLVIPVTMGFYVYKVNALAC 589  
DB 509 PFLMEKPLRPGSSLGLPFASYSFFVIRNAKVAAC 542

RESULT 11  
US-09-435-739-10  
Sequence 10, Application US/09435739  
Patent No. 6664105

GENERAL INFORMATION:  
APPLICANT: Pecker, Iris  
APPLICANT: Vlodavsky, Israel  
APPLICANT: Feinstein, Elena  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
FILE REFERENCE: 00/20454  
CURRENT APPLICATION NUMBER: US/09/435,739  
CURRENT FILING DATE: 2001-06-05  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 10  
LENGTH: 543  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-435-739-10

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

18 PPACIAGALYIALLLHLSSQAGDRRLPVDRRAAGLKEKTLILDVSTKRVRYTEN 77  
DB PLGLSPGAL-----PRPA-----QAQDVVDLDFFTQEPHLVSPS 55  
QY FLSLQDPSIITH-GWLDPLSSKRLVTLARGLSPAFIRFGGKRTDFLOFQNLNPAKSRG 136  
DB FLSTVIDANLATDPRFLILGSPKLRRLARGLSPAYIRFGGKTDFLIF---DPKEST 111  
QY 137 GPGPDYLYKNVEDIVRSQVADLQKQCKIAQ-HPDVMVLQREKAQOMELVLLKEQFSN 195  
DB 112 FEERSYMQSQVNODI-----CKYGSIPPDVEEKLRLMPYQEQ-LIREHYQK 158  
QY 196 TYSNLIILFASLIDKLYNFADCGSLHIFALNALRRNNNNSSSSALSLKYSASKKYN 255  
DB 159 KFKNSTYRSRSDVLYITFANCSGDLIFGLNALRTADLQWSSNAQLLDYCCKSGYNI 218  
QY 256 SWEIGNEPNRYTMMGRAVNGSQLGKDYIQLKSLQPIRTYRSALYGPNIQPRKXVIA 315  
DB 219 SWEIGNEPNRYTMMGRAVNGSQLGKDYIQLKSLQPIRTYRSALYGPNIQPRKXVIA 315  
QY 316 LLDGFMKVAAGSTVDAVTWQHCVYIDGRVYKVMDFLKTLLDLSQIRKIQKVNTYTPGK 375  
DB 278 MLKSTFKAGGEVIDSVTHHYLYNKRTRATREDFLNPVDLIFISSQKVPQVSESTRPK 337  
QY 376 KIMLEGVVTTSAGTNNISDSYAAGFLMNTLGMLANOGIDVIRHSFFDHGYNHLYDON 435  
DB 338 KWLGETSSAYGGAPLSDITFAAGFMWLDKLGSAEMGIEVMQVFFGAGNTHLVDEN 397  
QY 436 ENPLPDYMLSLYKLLIGPKYLAHVAGLQKRPGRVIRDKLRIYACNTHNNHNYVVG 495  
DB 398 FDPDPYMLSLFVKLVGTQVLMASVQSKRR-----KLRYVLCNTNDNRYKEG 448  
QY 496 SITFLIINLHRSRKKIKLAGTRDLYVHGYLLQPYGQELSKSVQVINGQPLVWDDGT 555  
DB 449 DLTVAIINLHNTKYLRLPYFPNSKQVQDKYLLRPLGPHGLSKSVQVINGQPLVWDDGT 508  
QY 556 PELKRPRLRAGRTLVIPVTMGFYVYKVNALAC 589  
DB 509 PFLMEKPLRPGSSLGLPFASYSFFVIRNAKVAAC 542

RESULT 12  
US-09-930-218-3  
Sequence 3, Application US/09930218  
Patent No. 6677137

GENERAL INFORMATION:  
APPLICANT: Goldsmith, Orit  
APPLICANT: Pecker, Iris  
APPLICANT: Vlodavsky, Israel  
APPLICANT: Israel, Michael  
TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H  
FILE REFERENCE: 01/22335  
CURRENT APPLICATION NUMBER: US/09/930,218  
CURRENT FILING DATE: 2001-08-16  
PRIOR APPLICATION NUMBER: 09/666,390  
PRIOR FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 543  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-930-218-3

Query Match 37.2%; Score 1148.5; DB 4; Length 543;  
Best Local Similarity 43.4%; Pred. No. 5.3e-110;  
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

18 PPACIAGALYIALLLHLSSQAGDRRLPVDRRAAGLKEKTLILDVSTKRVRYTEN 77  
DB 20 PLGLSPGAL-----PRPA-----QAQDVVDLDFFTQEPHLVSPS 55

```

QY 78 FLSLQDPSIIHD-GWLDPLSSKRLVTLARGLSPALFRFGKRTDPLQFONLRNPAKSRG 136
DB 56 FLSVITDANLADPRLILILGSPKRLTLARGLSPALFRFGKRTDPLIF-----DPKKST 111
QY 137 GPEPDYLLKNVEDDIYRSVDALDKQKCKIAQ-HPDVMVLVOREKAQMHVLLKEQFSN 195
DB 112 FEERSWQSQVNODI-----CKYGSIPDVEEKLRLMEWYEQOL-LLEHRYOK 158
QY 196 TYSNLIITARSJDKLVNFDSCGLHIFALNALRRPNNSWSSSALSLKTSASKKXNI 255
DB 159 KFNKSTYSSRSVDLYTFANCSGLDLIFGLNALRTRADLOMNSNQLLLDYCSSKXNI 218
QY 256 SWEIGNEPNRRTMGRANVGSOLGKDYIQLKSLQPIRYSASLYGNIGRPRKNVIA 315
DB 219 SWEIGNEPNRFLKKADIFINGSQLGEDIYQLHKLKLR-STFKNAKLYGPDVQPRKRTAK 277
QY 316 LLDGFMKVASSTVDATWQHCYIDGRVYKMDLKTLLDLSQIRKIQKVTNTYTPGK 375
DB 278 MKKSPFKAGGEVIDSVTWHYINGRATREDPLNDVDLIFISSVQKVFQVVESTRPGK 337
QY 376 KIMLEGVTTSSAGTNNLSYSYAGFLMNTLGMANOGIDVYIRSFDPHGYNHLVDON 435
DB 338 KVMLEGTSSAYGGAPILSDTFAGFWMLDKLGISARMGLEVVRQVFFAGNYHLVDEN 397
QY 436 FNPJPDYWLSTLYKRLIGPKYLAHVAVAGLQKRPKRVIRDKLRIYAHCTNNHNNHYRG 495
DB 398 FDPJPDYWLSTLIFKXLVGTKVLASVQGSKR-----KLRYVLIHCTNTDNPYKKG 448
QY 496 SITFIINLHRSRKKIKLAGTLRDKLVHGYLLOPYGQEGLSKSVQNLGCPVMVDGTL 555
DB 449 DLTLYALNHVNTKYTLRPPFSNKQYDKILRLRPLGPHGLSSVQNLGTLKXVDQTL 508
QY 556 PELKRPRLRAGRTLVIPVTMGFYVKNVNAALAC 589
DB 509 PLMEKPLRPGSSILGPAFSYSFVIRNNAKVAAC 542

RESULT 13
US-09-435-739-14
; Sequence 14, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Becker, Iris
; APPLICANT: Viodevsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435,739
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-435-739-14

Query Match 37.2%; Score 1148.5; DB 4; Length 592;
Best Local Similarity 43.4%; Pred. No. 6.1e-110;
Matches 249; Conservative 82; Mismatches 190; Indels 53; Gaps 9;

```

```

QY 18 PPACIAPGALYLLALHLSLSQAGDRRLPVDRRAAGKEKTLILLDVSTKRPVRYTEN 77
DB 69 PLGPISPGAL-----PRPA-----QAQVDVLDLFFTEQEPILHVSPS 104
QY 78 FLSLQDPSIIHD-GWLDPLSSKRLVTLARGLSPALFRFGKRTDPLQFONLRNPAKSRG 136
DB 105 FLSVITDANLADPRLILILGSPKRLTLARGLSPALFRFGKRTDPLIF-----DPKKST 160
QY 137 GPEPDYLLKNVEDDIYRSVDALDKQKCKIAQ-HPDVMVLVOREKAQMHVLLKEQFSN 195

```

```

DB 161 FEERSWQSQVNODI-----CKYGSIPDVEEKLRLMEWYEQOL-LLEHRYOK 207
QY 196 TYSNLIITARSJDKLVNFDSCGLHIFALNALRRPNNSWSSSALSLKTSASKKXNI 255
DB 208 KFNKSTYSSRSVDLYTFANCSGLDLIFGLNALRTRADLOMNSNQLLLDYCSSKXNI 267
QY 256 SWEIGNEPNRRTMGRANVGSOLGKDYIQLKSLQPIRYSASLYGNIGRPRKNVIA 315
DB 268 SWEIGNEPNRFLKKADIFINGSQLGEDIYQLHKLKLR-STFKNAKLYGPDVQPRKRTAK 326
QY 316 LLDGFMKVASSTVDATWQHCYIDGRVYKMDLKTLLDLSQIRKIQKVTNTYTPGK 375
DB 327 MKKSPFKAGGEVIDSVTWHYINGRATREDPLNDVDLIFISSVQKVFQVVESTRPGK 366
QY 376 KIMLEGVTTSSAGTNNLSYSYAGFLMNTLGMANOGIDVYIRSFDPHGYNHLVDON 435
DB 387 KVMLEGTSSAYGGAPILSDTFAGFWMLDKLGISARMGLEVVRQVFFAGNYHLVDEN 446
QY 436 FNPJPDYWLSTLYKRLIGPKYLAHVAVAGLQKRPKRVIRDKLRIYAHCTNNHNNHYRG 495
DB 447 FDPJPDYWLSTLIFKXLVGTKVLASVQGSKR-----KLRYVLIHCTNTDNPYKKG 497
QY 496 SITFIINLHRSRKKIKLAGTLRDKLVHGYLLOPYGQEGLSKSVQNLGCPVMVDGTL 555
DB 498 DLTLYALNHVNTKYTLRPPFSNKQYDKILRLRPLGPHGLSSVQNLGTLKXVDQTL 557
QY 556 PELKRPRLRAGRTLVIPVTMGFYVKNVNAALAC 589
DB 558 PLMEKPLRPGSSILGPAFSYSFVIRNNAKVAAC 591

RESULT 14
US-09-181-336-13
; Sequence 13, Application US/09181336A
; Patent No. 6242238
; GENERAL INFORMATION:
; APPLICANT: FREEMAN, Craig Geoffrey
; APPLICANT: PARISH, Richard
; APPLICANT: HAMDORF, Brenton James
; APPLICANT: HOLETT, Mark Darren
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING MAMMALIAN
; TITLE OF INVENTION: ENDONUCLEONIDASE AND USES THEREFOR
; FILE REFERENCE: 032505-004
; CURRENT APPLICATION NUMBER: US/09/181,336A
; CURRENT FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: AU PP0062
; EARLIER FILING DATE: 1997-10-28
; EARLIER APPLICATION NUMBER: AU PP0812
; EARLIER FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-336-13

Query Match 37.1%; Score 1144.5; DB 3; Length 543;
Best Local Similarity 43.2%; Pred. No. 1.4e-109;
Matches 248; Conservative 83; Mismatches 190; Indels 53; Gaps 9;

```

```

QY 18 PPACIAPGALYLLALHLSLSQAGDRRLPVDRRAAGKEKTLILLDVSTKRPVRYTEN 77
DB 20 PLGPISPGAL-----PRPA-----QAQVDVLDLFFTEQEPILHVSPS 55
QY 78 FLSLQDPSIIHD-GWLDPLSSKRLVTLARGLSPALFRFGKRTDPLQFONLRNPAKSRG 136
DB 56 FLSVITDANLADPRLILILGSPKRLTLARGLSPALFRFGKRTDPLIF-----DPKKST 111
QY 137 GPEPDYLLKNVEDDIYRSVDALDKQKCKIAQ-HPDVMVLVOREKAQMHVLLKEQFSN 195
DB 112 FEERSWQSQVNODI-----CKYGSIPDVEEKLRLMEWYEQOL-LLEHRYOK 158

```

```

QY 196 TYSNLTARSLDKLYNEADCGSLHILFALNLRPNNSWSSSALSLKYSASKYNI 255
DB 159 KFNKSTYSRSSVDVLYTFANCSGLDIFGLNALRLRADLQNMSSNQLLLDCCSKGYNI 218
QY 256 SWEIGNEPNRYTRMGRANVAGSOLGKDYIOLKSLLOPIRYSRASYGNIGRPRKNVIA 315
DB 219 SWEIGNEPNRSFLKKADIFINGSQLGDFIOLHKLARK-STFKNAKLYGPDVQGPBRKTAK 277
QY 316 LLDGEMKVASGSTDVATWQHCTYDGRVVKYMPFLKTRLLDLSOQIRKIQKVVNTYTPPK 375
DB 278 MKSFLKAGGEVIDSVTMHHYILNGRTAREDFLNPDVLDIFISSVQKVFQVVESTRPK 337
QY 376 KIMLEGVTTTSAGCTNNLSDSYAAGFLMNTLQMLANOGIDVYIRHSFPHGYNHLYDON 435
DB 338 KVMLEGETSSAYGGAPLSDTFPAAGFMWLDKLSARMGIEVMRQVFFGAGNYHLVDEN 397
QY 436 FNEPLDYWLSLLYKRLIGPKVLAHVHAGLQKRPGRVIRDKLRIYACHTNNHNNYVRG 495
DB 398 FDLPLDYWLSLLYKRLIGPKVLAHVHAGLQKRPGRVIRDKLRIYACHTNNHNNYVRG 448
QY 496 SITFLFINLHRSRKIKLAGTLADKLYHOYLLQPYGQEGKSKSVQNLNGOPLVMVDGTL 555
DB 449 DLTLYAINLHNVTKYLRLLPYFSSNKQVDKYLRLPLGPHGLSSVQNLNGLTLKAVDQTL 508
QY 556 PELKPRPLRAGRTLVIPTVMGFYVKNVNALAC 589
DB 509 PLMEKPLRPGSSSLGLPAPFSYFVIRNAKVAAC 542

```

```

RESULT 15
US-09-601-777-2
/ Sequence 2, Application US/09601777
/ Patent No. 6461848
/ GENERAL INFORMATION:
/ APPLICANT: Nakajima, Motowo
/ APPLICANT: Funakado, Minako
/ TITLE OF INVENTION: Human heparanase polypeptide and cDNA
/ FILE REFERENCE: 30384A
/ CURRENT APPLICATION NUMBER: US/09/601,777
/ CURRENT FILING DATE: 2000-08-07
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: Patentln Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 588
/ TYPE: PRT
/ ORGANISM: Human
US-09-601-777-2

```

```

Query Match 37.1%; Score 1144.5; DB 4; Length 588;
Best Local Similarity 43.2%; Pred. No. 1.6e-109; Indels 53; Gaps 9;
Matches 248; Conservative 83; Mismatches 190;

```

```

QY 18 PPACLAAGALYLLALLHLSSLSQAGRRPLVDRAAGLKEKTLILLDVSTKNPRTVNNEN 77
DB 65 FLGLSPGAL-----PRPA-----QAQVVDLDFQEBLHVSPS 100
QY 78 FLISQLDPSIHD-GWIDFLSKRLVTLAGLSAPFLRFGKRTDFLQFQNLNPAKSRG 136
DB 101 FLSTYTDANLATDRFRLILGSPRLRLARGLSPAYRFGRTKDFLIF---DPKEEST 156
QY 137 GPGEDVYLYKNVEDDIVASDALDKOKGCKIAQ-HPDVMVLQREKAAQMLVLLKEQFSN 195
DB 157 FEERSYVQSQVNDI-----CKGSIIPDVEEKLRLWEPYQEQTL-LLEHYQK 203
QY 196 TYSNLTITASLDKLYNFADCSGLHILFALNALRPNNSWSSSALSLKYSASKYNI 255
DB 204 KFNKSTYSRSSVDVLYTFANCSGLDIFGLNALRLRADLQNMSSNQLLLDCCSKGYNI 263
QY 256 SWEIGNEPNRYTRMGRANVAGSOLGKDYIOLKSLLOPIRYSRASYGNIGRPRKNVIA 315
DB 264 SWEIGNEPNRSFLKKADIFINGSQLGDFIOLHKLARK-STFKNAKLYGPDVQGPBRKTAK 322
QY 316 LLDGEMKVASGSTDVATWQHCTYDGRVVKYMPFLKTRLLDLSOQIRKIQKVVNTYTPPK 375

```

```

DB 323 MKSFLKAGGEVIDSVTMHHYILNGRTAREDFLNPDVLDIFISSVQKVFQVVESTRPKG 382
QY 376 KIMLEGVTTTSAGCTNNLSDSYAAGFLMNTLQMLANOGIDVYIRHSFPHGYNHLYDON 435
DB 383 KVMLEGETSSAYGGAPLSDTFPAAGFMWLDKLSARMGIEVMRQVFFGAGNYHLVDEN 442
QY 436 FNEPLDYWLSLLYKRLIGPKVLAHVHAGLQKRPGRVIRDKLRIYACHTNNHNNYVRG 495
DB 443 FDLPLDYWLSLLYKRLIGPKVLAHVHAGLQKRPGRVIRDKLRIYACHTNNHNNYVRG 493
QY 496 SITFLFINLHRSRKIKLAGTLADKLYHOYLLQPYGQEGKSKSVQNLNGOPLVMVDGTL 555
DB 494 DLTLYAINLHNVTKYLRLLPYFSSNKQVDKYLRLPLGPHGLSSVQNLNGLTLKAVDQTL 553
QY 556 PELKPRPLRAGRTLVIPTVMGFYVKNVNALAC 589
DB 554 PLMEKPLRPGSSSLGLPAPFSYFVIRNAKVAAC 587

```

```

Search completed: May 6, 2004, 13:46:48
Job time : 24 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:45:47 ; Search time 49 Seconds  
(without alignments)  
3353.460 Million cell updates/sec

Title: US-10-088-676-2  
Perfect score: 3088  
Sequence: 1 MRVLCAPPEAMPSSNSRPPA.....FTVMGFYVKNVNALACRYR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 27756755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3078	99.7	592	US-09-978-249-13	Sequence 13, Appl
2	3071	99.4	592	US-10-177-245A-2	Sequence 2, Appl
3	2740	88.7	534	US-09-880-262-2	Sequence 2, Appl
4	2736	88.6	534	US-09-836-461-2	Sequence 2, Appl
5	2736	88.6	534	US-10-177-245A-4	Sequence 4, Appl
6	2446	79.2	480	US-09-978-249-7	Sequence 7, Appl
7	2446	79.2	480	US-10-177-245A-6	Sequence 6, Appl
8	2238	72.5	439	US-09-978-249-12	Sequence 12, Appl
9	2238	72.5	439	US-09-978-249-12	Sequence 12, Appl
10	1165	37.7	536	US-10-431-438-2	Sequence 2, Appl
11	1163.5	37.7	535	US-09-930-218-1	Sequence 1, Appl
12	1163.5	37.7	535	US-09-776-874A-44	Sequence 44, Appl
13	1163.5	37.7	535	US-09-988-113-44	Sequence 44, Appl
14	1163.5	37.7	535	US-10-341-582-44	Sequence 44, Appl
15	1163.5	37.7	535	US-10-384-451-44	Sequence 44, Appl

16	1163.5	37.7	535	US-10-431-438-1	Sequence 1, Appl
17	1163.5	37.7	535	US-10-384-450-44	Sequence 44, Appl
18	1163.5	37.7	535	US-10-371-218A-44	Sequence 44, Appl
19	1163.5	37.7	535	US-10-456-573-44	Sequence 44, Appl
20	1149	37.2	545	US-09-899-440-18	Sequence 18, Appl
21	1149	37.2	545	US-10-115-479-42	Sequence 42, Appl
22	1148.5	37.2	543	US-09-759-207-2	Sequence 2, Appl
23	1148.5	37.2	543	US-09-930-218-3	Sequence 3, Appl
24	1148.5	37.2	543	US-09-186-200-1	Sequence 1, Appl
25	1148.5	37.2	543	US-09-944-602-2	Sequence 2, Appl
26	1148.5	37.2	543	US-09-322-977-2	Sequence 2, Appl
27	1148.5	37.2	543	US-09-988-113-10	Sequence 10, Appl
28	1148.5	37.2	543	US-10-676-079-2	Sequence 2, Appl
29	1148.5	37.2	543	US-10-137-351-2	Sequence 2, Appl
30	1148.5	37.2	543	US-10-341-582-10	Sequence 10, Appl
31	1148.5	37.2	543	US-10-384-451-10	Sequence 10, Appl
32	1148.5	37.2	543	US-10-431-438-3	Sequence 3, Appl
33	1148.5	37.2	543	US-10-368-044A-1	Sequence 1, Appl
34	1148.5	37.2	543	US-10-384-450-10	Sequence 10, Appl
35	1148.5	37.2	543	US-10-371-218A-10	Sequence 10, Appl
36	1148.5	37.2	543	US-10-456-573-10	Sequence 10, Appl
37	1148.5	37.2	543	US-09-776-874A-14	Sequence 14, Appl
38	1148.5	37.2	543	US-09-988-113-14	Sequence 14, Appl
39	1148.5	37.2	543	US-10-341-582-14	Sequence 14, Appl
40	1148.5	37.2	543	US-10-384-451-14	Sequence 14, Appl
41	1148.5	37.2	543	US-10-384-450-14	Sequence 14, Appl
42	1148.5	37.2	543	US-10-371-218A-14	Sequence 14, Appl
43	1148.5	37.2	543	US-10-456-573-14	Sequence 14, Appl
44	1148.5	37.2	543	US-10-314-683-2	Sequence 2, Appl
45	1144.5	37.1	554	US-10-314-683-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-978-249-13  
Sequence 13, Application US/09978249  
Patent No. US20020106780A1  
GENERAL INFORMATION:  
APPLICANT: Fisiella, et al.  
TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies  
FILE REFERENCE: PTO5491  
CURRENT FILING DATE: 2001-10-17  
CURRENT FILING DATE: 2001-10-17  
PRIORITY FILING DATE: 2001-04-11  
PRIORITY FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: PCT/US01/11643  
PRIOR APPLICATION NUMBER: 60/198,123  
PRIOR FILING DATE: 2000-04-18  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 13  
LENGTH: 592  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-978-249-13

Query Match	99.7%	Score 3078;	DB 9;	Length 592;
Best Local Similarity	99.7%	Pred. No. 4.3e-25;		
Matches 590;	Conservative	1;	Mismatches	1;
			Indels	0;
			Gaps	0;
QY	1	MRVLCAPPEAMPSSNSRPPACLAGALYLLALHLSSSQGDRPPVPDAAAGLKEKTL	60	
DB	1	MRVLCAPPEAMPSSNSRPPACLAGALYLLALHLSSSQGDRPPVPDAAAGLKEKTL	60	
QY	61	ILDVSTKNPVTVNENFLSLQDPSTIHDGMLDFLSKRLVTLARGSPAFLEFGKRT	120	
DB	61	ILDVSTKNPVTVNENFLSLQDPSTIHDGMLDFLSKRLVTLARGSPAFLEFGKRT	120	
QY	121	DLQFQNTLRNPAKSGGPGPDYLLKNYEDD1VRSVADLKQKCKIAQHPVMTLQREK	180	
DB	121	DLQFQNTLRNPAKSGGPGPDYLLKNYEDD1VRSVADLKQKCKIAQHPVMTLQREK	180	

```

QY 181 AACMHLVLLKEQPSNTYSNLLTARSLDKYNFADCSGHLIPALNALRRPNNSNSS 240
DB 181 AACMHLVLLKEQPSNTYSNLLTARSLDKYNFADCSGHLIPALNALRRPNNSNSS 240
QY 241 ALSILKYSASAKKNIISWELGNEPNRYRTWGRAVNSQOLKDYIOUKSLQPIRISRAS 300
DB 241 ALSILKYSASAKKNIISWELGNEPNRYRTWGRAVNSQOLKDYIOUKSLQPIRISRAS 300
QY 301 LYGPNI GRPKNYTALIDGFMKTAGSTVDVAVTWOHCYIDGRVVKWDFKTRLLDLSQ 360
DB 301 LYGPNI GRPKNYTALIDGFMKTAGSTVDVAVTWOHCYIDGRVVKWDFKTRLLDLSQ 360
QY 361 IRIKQVNTYTPGKIMLEGVTTSGAGTNNLSDSYAAGFLMNTLGMLANQIDVIR 420
DB 361 IRIKQVNTYTPGKIMLEGVTTSGAGTNNLSDSYAAGFLMNTLGMLANQIDVIR 420
QY 421 HSFPDHGYNHLDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKRI 480
DB 421 HSFPDHGYNHLDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKRI 480
QY 481 YAHCTNNHNNHYVRSITLFIINLHRSRKKIKLAGTRDLVHQYLLQPYGQGLKSKSV 540
DB 481 YAHCTNNHNNHYVRSITLFIINLHRSRKKIKLAGTRDLVHQYLLQPYGQGLKSKSV 540
QY 541 QLNQGPLVMVDGTLPELKRPPLAGRTLVIPTVTMGFFVVKVNNALACRYR 592
DB 541 QLNQGPLVMVDGTLPELKRPPLAGRTLVIPTVTMGFFVVKVNNALACRYR 592

```

RESULT 2  
US-10-177-245A-2  
Sequence 2, Application US/10177245A  
Publication No. US20030083254A1

```

GENERAL INFORMATION:
APPLICANT: McKenzie, Edward Alexander
APPLICANT: Stamps, Alasdair Craig
APPLICANT: Terrett, Jonathan Alexander
APPLICANT: Tyson, Kerry Louise
TITLE OF INVENTION: Substances
FILE REFERENCE: 2543-1-027
CURRENT APPLICATION NUMBER: US/10/177,245A
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: PCT/GB00/04963
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: UK 0008713.0
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: UK 9930392.7
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 592
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(592)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-177-245A-2

```

```

Query Match 99.4%; Score 3071; DB 14; Length 592;
Best Local Similarity 99.5%; Pred. No. 2, 1e-294;
Matches 589; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MRVLCAPFAPMSSNSRPPACIAPGALYIALLLHLSISSQAGDRRPLFVDBRAAGLKEKTL 60
DB 1 MRVLCAPFAPMSSNSRPPACIAPGALYIALLLHLSISSQAGDRRPLFVDBRAAGLKEKTL 60
QY 61 ILIDVSTKNPVTNENFLSLQDPSIHDGMDPLSSKRLVTLARGISPAFLRGGKRT 120
DB 61 ILIDVSTKNPVTNENFLSLQDPSIHDGMDPLSSKRLVTLARGISPAFLRGGKRT 120
QY 121 DFLQFQNLNRPKASRGSGPDPYLLKNYEDDIVRSVDALDKQKGIAGHPDVMVLQREK 180
DB 121 DFLQFQNLNRPKASRGSGPDPYLLKNYEDDIVRSVDALDKQKGIAGHPDVMVLQREK 180

```

```

DB 121 DFLQFQNLNRPKASRGSGPDPYLLKNYEDDIVRSVDALDKQKGIAGHPDVMVLQREK 180
QY 181 AACMHLVLLKEQPSNTYSNLLTARSLDKYNFADCSGHLIPALNALRRPNNSNSS 240
DB 181 AACMHLVLLKEQPSNTYSNLLTARSLDKYNFADCSGHLIPALNALRRPNNSNSS 240
QY 241 ALSILKYSASAKKNIISWELGNEPNRYRTWGRAVNSQOLKDYIOUKSLQPIRISRAS 300
DB 241 ALSILKYSASAKKNIISWELGNEPNRYRTWGRAVNSQOLKDYIOUKSLQPIRISRAS 300
QY 301 LYGPNI GRPKNYTALIDGFMKTAGSTVDVAVTWOHCYIDGRVVKWDFKTRLLDLSQ 360
DB 301 LYGPNI GRPKNYTALIDGFMKTAGSTVDVAVTWOHCYIDGRVVKWDFKTRLLDLSQ 360
QY 361 IRIKQVNTYTPGKIMLEGVTTSGAGTNNLSDSYAAGFLMNTLGMLANQIDVIR 420
DB 361 IRIKQVNTYTPGKIMLEGVTTSGAGTNNLSDSYAAGFLMNTLGMLANQIDVIR 420
QY 421 HSFPDHGYNHLDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKRI 480
DB 421 HSFPDHGYNHLDONENPLDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKRI 480
QY 481 YAHCTNNHNNHYVRSITLFIINLHRSRKKIKLAGTRDLVHQYLLQPYGQGLKSKSV 540
DB 481 YAHCTNNHNNHYVRSITLFIINLHRSRKKIKLAGTRDLVHQYLLQPYGQGLKSKSV 540
QY 541 QLNQGPLVMVDGTLPELKRPPLAGRTLVIPTVTMGFFVVKVNNALACRYR 592
DB 541 QLNQGPLVMVDGTLPELKRPPLAGRTLVIPTVTMGFFVVKVNNALACRYR 592

```

RESULT 3  
US-09-880-262-2  
Sequence 2, Application US/09880262  
Patent No. US20020137907A1

```

GENERAL INFORMATION:
APPLICANT: LAMWINE, CHRISTOPHER DONALD
APPLICANT: SOUTHAIR, CHRISTOPHER GEOFFREY CARSON
APPLICANT: RANCE, KIM
APPLICANT: HAYES, PHILIP DAVID
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30231
CURRENT APPLICATION NUMBER: US/09/880,262
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: UK 0014447.7
PRIOR FILING DATE: 2000-06-13
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 534
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-880-262-2

```

```

Query Match 88.7%; Score 2740; DB 9; Length 534;
Best Local Similarity 90.0%; Pred. No. 1e-261;
Matches 533; Conservative 0; Mismatches 1; Indels 58; Gaps 1;
QY 1 MRVLCAPFAPMSSNSRPPACIAPGALYIALLLHLSISSQAGDRRPLFVDBRAAGLKEKTL 60
DB 1 MRVLCAPFAPMSSNSRPPACIAPGALYIALLLHLSISSQAGDRRPLFVDBRAAGLKEKTL 60
QY 61 ILIDVSTKNPVTNENFLSLQDPSIHDGMDPLSSKRLVTLARGISPAFLRGGKRT 120
DB 61 ILIDVSTKNPVTNENFLSLQDPSIHDGMDPLSSKRLVTLARGISPAFLRGGKRT 120
QY 121 DFLQFQNLNRPKASRGSGPDPYLLKNYEDDIVRSVDALDKQKGIAGHPDVMVLQREK 180
DB 121 DFLQFQNLNRPKASRGSGPDPYLLKNYEDDIVRSVDALDKQKGIAGHPDVMVLQREK 180
QY 181 AACMHLVLLKEQPSNTYSNLLTARSLDKYNFADCSGHLIPALNALRRPNNSNSS 240
DB 181 AACMHLVLLKEQPSNTYSNLLTARSLDKYNFADCSGHLIPALNALRRPNNSNSS 240

```



```

Db      181 AAQMHVLVLEKEQFSNTYSNLIIT----- 203
Qy      241 ALSILKYSASKKYNISWELGNEPNRYRTMHWGRAVNSQOLGKDYIOLKSLLOPIRISRAS 300
Db      204 -----EPNNRYRTMHWGRAVNSQOLGKDYIOLKSLLOPIRISRAS 242
Qy      301 LYGNIGRPRKNVIALLDGFMKVAAGSTVDVATWQHCHTIDGRVYKWMDFLKTRLDLDLSDQ 360
Db      243 LYGNIGRPRKNVIALLDGFMKVAAGSTVDVATWQHCHTIDGRVYKWMDFLKTRLDLDLSDQ 302
Qy      361 IRKIOKVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 420
Db      303 IRKIOKVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 362
Qy      421 HSFPDHGYNHLDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db      363 HSFPDHGYNHLDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 422
Qy      481 YAHCTNHNNHYVRSITLFTIINLHRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 540
Db      423 YAHCTNHNNHYVRSITLFTIINLHRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 482
Qy      541 QLNQGPLVWVDGTLPELKPRPLRAGRTLVIPVTMGFFVYKNNALACRYR 592
Db      483 QLNQGPLVWVDGTLPELKPRPLRAGRTLVIPVTMGFFVYKNNALACRYR 534

RESULT 4
US-09-836-461-2
; Sequence 2, Application US/09836461
; Patent No. US20020064853A1
; GENERAL INFORMATION:
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; TITLE OF INVENTION: Heparanase II, A No. US20020064853A1e1 Human Heparanase Paralog
; FILE REFERENCE: heparanase II
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-461-2

Query Match      88.6%; Score 2736; DB 9; Length 534;
Best Local Similarity 89.9%; Pred. No. 2,6e-261;
Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

Qy      1 MRVLCAPPEAMPSSNSRPPACIAPGALYALLLHLSSSQAGDRRPLPVDRAGLKEKTL 60
Db      1 MRVLCAPPEAMPSSNSRPPACIAPGALYALLLHLSSSQAGDRRPLPVDRAGLKEKTL 60
Qy      61 ILDVDSTKNPRTVNTENFISQLDPSIITHDGLDPLSKRLVTLARGLSPAFLRFGGKRT 120
Db      61 ILDVDSTKNPRTVNTENFISQLDPSIITHDGLDPLSKRLVTLARGLSPAFLRFGGKRT 120
Qy      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDLADKQKGTIAQHPDVMVLQREK 180
Db      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDLADKQKGTIAQHPDVMVLQREK 180
Qy      181 AAQMHVLVLEKEQFSNTYSNLIITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Db      181 AAQMHVLVLEKEQFSNTYSNLIITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Qy      241 ALSILKYSASKKYNISWELGNEPNRYRTMHWGRAVNSQOLGKDYIOLKSLLOPIRISRAS 300
Db      204 -----EPNNRYRTMHWGRAVNSQOLGKDYIOLKSLLOPIRISRAS 242
Qy      301 LYGNIGRPRKNVIALLDGFMKVAAGSTVDVATWQHCHTIDGRVYKWMDFLKTRLDLDLSDQ 360
Db      243 LYGNIGRPRKNVIALLDGFMKVAAGSTVDVATWQHCHTIDGRVYKWMDFLKTRLDLDLSDQ 302

```

```

Qy      361 IRKIOKVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 420
Db      303 IRKIOKVNTYTPGKKTMLBEGVTTTSAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 362
Qy      421 HSFPDHGYNHLDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db      363 HSFPDHGYNHLDQNFNLPDYWLSLYKRLIGKVLAVHAGLQKRPGRVIRDKLRI 422
Qy      481 YAHCTNHNNHYVRSITLFTIINLHRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 540
Db      423 YAHCTNHNNHYVRSITLFTIINLHRSRKKIKLAGTLDKLVHGYLLQPYQOGLKSKSV 482
Qy      541 QLNQGPLVWVDGTLPELKPRPLRAGRTLVIPVTMGFFVYKNNALACRYR 592
Db      483 QLNQGPLVWVDGTLPELKPRPLRAGRTLVIPVTMGFFVYKNNALACRYR 534

```

```

RESULT 5
US-10-177-245A-4
; Sequence 4, Application US/10177245A
; Publication No. US20030083254A1
; GENERAL INFORMATION:
; APPLICANT: McKenzie, Edward Alexander
; APPLICANT: Stamps, Alasdair Craig
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Substances
; FILE REFERENCE: 2543-1-027
; CURRENT APPLICATION NUMBER: US/10/177,245A
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: PCT/GB00/04963
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: UK 0008713.0
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: UK 9930392.7
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-245A-4

```

```

Query Match      88.6%; Score 2736; DB 14; Length 534;
Best Local Similarity 89.9%; Pred. No. 2,6e-261;
Matches 532; Conservative 1; Mismatches 1; Indels 58; Gaps 1;

Qy      1 MRVLCAPPEAMPSSNSRPPACIAPGALYALLLHLSSSQAGDRRPLPVDRAGLKEKTL 60
Db      1 MRVLCAPPEAMPSSNSRPPACIAPGALYALLLHLSSSQAGDRRPLPVDRAGLKEKTL 60
Qy      61 ILDVDSTKNPRTVNTENFISQLDPSIITHDGLDPLSKRLVTLARGLSPAFLRFGGKRT 120
Db      61 ILDVDSTKNPRTVNTENFISQLDPSIITHDGLDPLSKRLVTLARGLSPAFLRFGGKRT 120
Qy      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDLADKQKGTIAQHPDVMVLQREK 180
Db      121 DFLQFONLRNPAKSRGPGPDYLLKNYEDDIYRSVDLADKQKGTIAQHPDVMVLQREK 180
Qy      181 AAQMHVLVLEKEQFSNTYSNLIITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Db      181 AAQMHVLVLEKEQFSNTYSNLIITARSCLKYNFADCSGLHLIFALNALRRPNNSWSSS 240
Qy      241 ALSILKYSASKKYNISWELGNEPNRYRTMHWGRAVNSQOLGKDYIOLKSLLOPIRISRAS 300
Db      204 -----EPNNRYRTMHWGRAVNSQOLGKDYIOLKSLLOPIRISRAS 242
Qy      301 LYGNIGRPRKNVIALLDGFMKVAAGSTVDVATWQHCHTIDGRVYKWMDFLKTRLDLDLSDQ 360
Db      243 LYGNIGRPRKNVIALLDGFMKVAAGSTVDVATWQHCHTIDGRVYKWMDFLKTRLDLDLSDQ 302

```

```

QY 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
Db 303 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 362
QY 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db 363 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 422
QY 481 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTRDKLVHGYLLQPYGEGLSKSV 540
Db 423 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTRDKLVHGYLLQPYGEGLSKSV 482
QY 541 QUNGQPLVMVDDGTLPELKRPPLRAGRTLVIPVTMGFYVKNVNALACRYR 592
Db 483 QUNGQPLVMVDDGTLPELKRPPLRAGRTLVIPVTMGFYVKNVNALACRYR 534

```

```

RESULT 6
US-09-978-249-7
; Sequence 7, Application US/09978249
; Patent No. US20020106780A1
; GENERAL INFORMATION:
; APPLICANT: Fiscella, et al.
; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P05441
; CURRENT APPLICATION NUMBER: US/09/978, 249
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/11643
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/198,123
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-249-7

```

```

Query Match 79.2%; Score 2446; DB 9; Length 480;
Best Local Similarity 80.9%; Pred. No. 1.1e-232;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

QY 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRAAGLKEKTL 60
Db 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRAAGLKEKTL 60
QY 61 ILIDVSTKNPRTVTVNENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
Db 61 ILIDVSTKNPRTVTVNENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
QY 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 180
Db 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 180
QY 149 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 149
Db 149 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 149
QY 181 AAQWHLVLLKEQFSNTYSNMLITARSCLKYFPADCSGLHLIFALNALRRPNNSWSSS 240
Db 181 AAQWHLVLLKEQFSNTYSNMLITARSCLKYFPADCSGLHLIFALNALRRPNNSWSSS 240
QY 149 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 149
Db 149 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 149
QY 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 300
Db 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 300
QY 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 188
Db 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 188
QY 301 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 360
Db 301 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 360
QY 189 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 248
Db 189 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 248
QY 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
Db 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
QY 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 480

```

```

Db 309 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 368
QY 481 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTRDKLVHGYLLQPYGEGLSKSV 540
Db 369 YAHCTNNHNNHYVGSITLFTINLHRSRKIKLAGTRDKLVHGYLLQPYGEGLSKSV 428
QY 541 QUNGQPLVMVDDGTLPELKRPPLRAGRTLVIPVTMGFYVKNVNALACRYR 592
Db 429 QUNGQPLVMVDDGTLPELKRPPLRAGRTLVIPVTMGFYVKNVNALACRYR 480

```

```

RESULT 7
US-10-177-245A-6
; Sequence 6, Application US/10177245A
; Publication No. US20030083254A1
; GENERAL INFORMATION:
; APPLICANT: McKenzie, Edward Alexander
; APPLICANT: Stamps, Alasdair Craig
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Substances
; FILE REFERENCE: 2543-1-027
; CURRENT APPLICATION NUMBER: US/10/177,245A
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: PCT/GB00/04963
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: UK 0008713.0
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: UK 9930392.7
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-245A-6

```

```

Query Match 79.2%; Score 2446; DB 14; Length 480;
Best Local Similarity 80.9%; Pred. No. 1.1e-232;
Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

QY 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRAAGLKEKTL 60
Db 1 MRVLCAPPEAMPSSNSRPPACIAPGALYIALHLHLSLSSQAGDRRPLVDRAAGLKEKTL 60
QY 61 ILIDVSTKNPRTVTVNENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
Db 61 ILIDVSTKNPRTVTVNENFLSLQDPSIHDGMLDFLSKRLVTLARGISPAFLRFGKRT 120
QY 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 180
Db 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 180
QY 149 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 149
Db 149 121 DFLQFQNLNPAKSGRGPDPYLLKNYEDIVRSVDALDKQCKIAQHPDVMVLQREK 149
QY 181 AAQWHLVLLKEQFSNTYSNMLITARSCLKYFPADCSGLHLIFALNALRRPNNSWSSS 240
Db 181 AAQWHLVLLKEQFSNTYSNMLITARSCLKYFPADCSGLHLIFALNALRRPNNSWSSS 240
QY 149 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 149
Db 149 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 149
QY 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 300
Db 241 ALSLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 300
QY 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 188
Db 150 -----EPNNYRTMHGRAVNGSOLGQDYIQLKSLQPIRIYSRAS 188
QY 301 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 360
Db 301 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 360
QY 189 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 248
Db 189 LYGPNIGRPKRVIALLDGFMKVAAGSTVDATVWQHCYIDGRVYKMDFLKTRLDITLSDQ 248
QY 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
Db 361 IRKIQKVNTYTTGGKIMLEGVTTTSAGTNNISDYAAGFLMNTLGMANGIDIVIR 420
QY 421 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 480
Db 309 HSFEDHGYNHLVDQNFNPLDPYMLSILYKELIGKVLAVHAGLQKRPGRVIRDKLRI 368

```

QY 481 YAHCTNHNHNHYRGSTITLFIINLHRSKKIKLAGTRDKLVHGYLLQPYGEGLSKSV 540  
 DB 369 YAHCTNHNHNHYRGSTITLFIINLHRSKKIKLAGTRDKLVHGYLLQPYGEGLSKSV 428  
 QY 541 QUNGOPLVMDGTLPELKRPRLRAGRTVIPPVTMGFYVKNVNLACRYR 592  
 DB 429 QUNGOPLVMDGTLPELKRPRLRAGRTVIPPVTMGFYVKNVNLACRYR 480

## RESULT 8

US-09-978-249-12  
 ; Sequence 12, Application US/09978249  
 ; Patent No. US20020106780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fiscella, et al.  
 ; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibodies  
 ; FILE REFERENCE: P054P1  
 ; CURRENT APPLICATION NUMBER: US/09/978,249  
 ; CURRENT FILING DATE: 2001-10-17  
 ; PRIOR APPLICATION NUMBER: PCT/US01/11643  
 ; PRIOR FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: 60/198,123  
 ; PRIOR FILING DATE: 2000-04-18  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 12  
 ; LENGTH: 439  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-978-249-12

Query Match 72.5%; Score 2238; DB 9; Length 439;  
 Best Local Similarity 79.5%; Pred. No. 3,66-212; Indels 112; Gaps 1;  
 Matches 438; Conservative 1; Mismatches 0;

QY 42 GDRRPLPVDRAGLKEKTLILLDVSTKNPVRTVNEFLSLQDPSIHDGMDLFLSKRL 101  
 DB 1 GDRRPLPVDRAGLKEKTLILLDVSTKNPVRTVNEFLSLQDPSIHDGMDLFLSKRL 60  
 QY 102 VTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEYDDIVRSDVALDKQ 161  
 DB 61 VTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEYDDIVRSDVALDKQ 108  
 QY 162 KCKGCKIAQHPVMLVLOREKAAQMHVLVILKEQFNTYSNLTITASLDKLYNFADCSGLH 221  
 DB 109 ----- 108  
 QY 222 IFALNALRPNPNNSWNSSSALSILKYSASKYINISWELGNEPNNTYRTMHGRAVNGSOLK 281  
 DB 109 ----- EPNNTYRTMHGRAVNGSOLK 128  
 QY 282 DYIOLKSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASITVDATWQHCHYIDGR 341  
 DB 129 DYIOLKSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASITVDATWQHCHYIDGR 188  
 QY 342 VVKYMDLFTKRLDLSQIRKIQVNTYTPGKKIMLEGVTTISAGTNNLSDSYAGF 401  
 DB 189 VVKYMDLFTKRLDLSQIRKIQVNTYTPGKKIMLEGVTTISAGTNNLSDSYAGF 248  
 QY 402 LMLNTLGMLANOGIDVIVIRHSFPDHGYNHLDQNFPLPDYWLISLYKRLIGPKYLAHV 461  
 DB 249 LMLNTLGMLANOGIDVIVIRHSFPDHGYNHLDQNFPLPDYWLISLYKRLIGPKYLAHV 308  
 QY 462 AGLOKRPGRVIVIDKLRITYAHCTNHNHNHYRGSTITLFIINLHRSKKIKLAGTRDKL 521  
 DB 309 AGLOKRPGRVIVIDKLRITYAHCTNHNHNHYRGSTITLFIINLHRSKKIKLAGTRDKL 368  
 QY 522 VHOYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPRLRAGRTVIPPVTMGFYV 581  
 DB 369 VHOYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPRLRAGRTVIPPVTMGFYV 428  
 QY 582 KVNNALACRYR 592

DB 429 KVNNALACRYR 439

## RESULT 9

US-09-930-218-2  
 ; Sequence 2, Application US/09930218  
 ; Patent No. US20020034810A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: goldsmith, orit  
 ; APPLICANT: pecker, iris  
 ; APPLICANT: vlodavsky, israel  
 ; TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H  
 ; TITLE OF INVENTION: HEPARANASE ACTIVITY  
 ; FILE REFERENCE: 01/22335  
 ; CURRENT APPLICATION NUMBER: US/09/930,218  
 ; CURRENT FILING DATE: 2001-08-16  
 ; PRIOR APPLICATION NUMBER: 09/666,390  
 ; PRIOR FILING DATE: 2000-09-20  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 536  
 ; TYPE: PRT  
 ; ORGANISM: Rattus rattus  
 ; US-09-930-218-2

Query Match 37.7%; Score 1165; DB 9; Length 536;  
 Best Local Similarity 45.1%; Pred. No. 7,56-106; Indels 28; Gaps 7;  
 Matches 248; Conservative 87; Mismatches 187;

QY 42 GDRRPLPVDRAGLKEKTLILLDVSTKNPVRTVNEFLSLQDPSIHDGMDLFLSKRL 99  
 DB 12 GRLRLTGTRPGTAPFTDQVLDLEFTRKLGQVSPSLSTIDASLATDRFLTGLSP 71  
 QY 100 RLVTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEYDDIVRSDVALD 159  
 DB 72 RLVTLARGLSPAFLRFGKRTDPLQFONLRNPAKSGGPGDYLYKNEYDDIVRSDVALD 123  
 QY 160 KCKGCKIAQHPVMLVLOREKAAQMHVLVILKEQFNTYSNLTITASLDKLYNFADCSGL 219  
 DB 124 -----RVSADVLKTLQWMPFQ-ELLILRQYREFRKNSTYSSSDMLYSFAKGR 175  
 QY 220 HIFALNALRPNPNNSWNSSSALSILKYSASKYINISWELGNEPNNTYRTMHGRAVNGSOL 279  
 DB 176 DLIFGLNALLRTPDRLRNSSNAQLLNYCSSKGYNISWELGNEPNNSFWKKAQISIDGLQ 235  
 QY 280 GKDYOQLSLQPIRIRYSASLYGNIGRPRKNVIALLDGFMKVASITVDATWQHCHYID 339  
 DB 236 GEDFVELHKLQK-SAFONAKLYGPDIGQPRKTYKILRSFLKAGGEVIDSLTWHYILN 294  
 QY 340 GEVYKMDLFTKRLDLSQIRKIQVNTYTPGKKIMLEGVTTISAGTNNLSDSYAA 399  
 DB 295 GRVATWEDEPLSDVLDLTFILSVOKILKYTKEMTPKKTKWLGTSAYGGAPLLSNTPAA 354  
 QY 400 GFLMNTLGMLANOGIDVIVIRHSFPDHGYNHLDQNFPLPDYWLISLYKRLIGPKYLAHV 459  
 DB 355 GFLMNTLGMLANOGIDVIVIRHSFPDHGYNHLDQNFPLPDYWLISLYKRLIGPKYLAHV 414  
 QY 460 HVAQLOKRPGRVIVIDKLRITYAHCTNHNHNHYRGSTITLFIINLHRSKKIKLAGTRDK 519  
 DB 415 RVKSGPD-----RSKLRVYIHCNTYVHPRYREGDILLYVANTLHNTYKHLKLPPEMFS 465  
 QY 520 KLVHGYILOPYGEGLSKSVQUNGOPLVMDGTLPELKRPRLRAGRTVIPPVTMGFYV 579  
 DB 466 RPDYKTLRPFSDGLSSVQUNGOPLVMDGTLPELKRPRLRAGRTVIPPVTMGFYV 525  
 QY 580 KVNNALACRYR 589  
 DB 526 VTRNAXIAAC 535

RESULT 10  
US-10-431-438-2  
Sequence 2, Application US/10431438  
Publication No. US20030180788A1  
GENERAL INFORMATION:  
APPLICANT: goldshmidt, orit  
APPLICANT: goldshmidt, orit  
APPLICANT: pecker, iris  
APPLICANT: vlodavsky, israel  
APPLICANT: israel, michael  
TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H  
TITLE OF INVENTION: HEPARANASE ACTIVITY  
FILE REFERENCE: 26013  
CURRENT APPLICATION NUMBER: US/10/431,438  
CURRENT FILING DATE: 2003-05-08  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 536  
TYPE: PRT  
ORGANISM: Rattus rattus  
US-10-431-438-2

Query Match 37.7%; Score 1165; DB 14; Length 536;  
Best Local Similarity 45.1%; Pred. No. 7.5e-106;  
Matches 248; Conservative 87; Mismatches 187; Indels 28; Gaps 7;  
QY 42 GDRRPLPYDRAAG-LKEKTLILDVSTKNPVYTNENFLSLQDPSIIHD-GWIDFLSSK 99  
DB 12 GLRALTLGTGTAAGTAFTDQVDVLEFYTKRLFGSVSPSLITIDASLATDPRFLTFGSGP 71  
QY 100 RLVTIARLSAPAFIRFGKRTDFOFQNLKRNPAKSGPGPDYLYKNYEDDIYRSVDVALD 159  
DB 72 RLRLARLSAPAFIRFGKRTDFOFQNLKRNPAKSGPGPDYLYKNYEDDIYRSVDVALD 123  
QY 160 KQKCKIAQHPDVMVLQREKRAQWHLVILKEQFSNTYSNLIITARSIDKLYNPAKSG 219  
DB 124 -----RVSADVLRKIQWMPFQ-ELLILREQYQREPKNSTYRSASVDMLYSFAKCSRL 175  
QY 220 HILFALNLRPNNSNNSSSALSLIKTSASKNTNISELGNERNYRTMGRVANSQ 279  
DB 176 DLIFLNLRLTLPDLRNNSNAQILLNCSKGNISWELNENPNSFWKKAQISIDQL 235  
QY 280 GKDYIOLKSLIOPRIYRSASLYGENIGRPKNVIALLDGFMKAGSTVDVAVTWOHCYID 339  
DB 236 GEDPEYELKILQK-SAFQNAKLYGPDIGQPKGVKILRSFLKAGGEVIDSLTWHHYLN 294  
QY 340 GRVAVKMDFLKRLDITLSQIRKIQKVNTYTPPKKIMLEGVVTTSAGGNNLSDSYAA 399  
DB 295 GRVAVKEDFLSDVDITFLISQKILKTKEMTPKKKWLGETSSAYGGGAPLLSNTFAA 354  
QY 400 GELMINTLQMLANOGIDIVIRHSPFDHGYNLVDONFNP.LPDYWL.SLLYKELIGPKYLA 459  
DB 335 GFMIDKLGSLAQGLIEVYMRQVFFGAGNYHLDENFRLPDYWL.SLLFKLVGPKYLA 414  
QY 460 HVAQLQRRPRGRVIRDKLRIYAHCTNNHNNYVGSITFLINLHRSRKIKLAGTLRD 519  
DB 415 RYKQPD-----RSKLRVYLHCTNYHPRYRGD.LLYVNLHNVTKILKPPRPF 465  
QY 520 KLVHGYLQPYGQGLKSKSVOLNGOPLVMVDGTLPELKRPPLRAGRTLVIPVTMGFY 579  
DB 466 RPYDKTLKLPFGSGDGLSKSVOLNGOPLKMDQTLPLTEKPLPAGSSLSVPASISYFF 525  
QY 580 VKNVNALAC 589  
DB 526 VLRNKAIAAC 535

RESULT 11  
US-09-930-218-1  
Sequence 1, Application US/09930218  
Patent No. US20020034810A1  
GENERAL INFORMATION:  
APPLICANT: goldshmidt, orit

APPLICANT: pecker, iris  
APPLICANT: vlodavsky, israel  
APPLICANT: israel, michael  
TITLE OF INVENTION: AVIAN AND REPTILE DERIVED POLYNUCLEOTIDE ENCODING A POLYPEPTIDE H  
TITLE OF INVENTION: HEPARANASE ACTIVITY  
FILE REFERENCE: 01/22335  
CURRENT APPLICATION NUMBER: US/09/930,218  
CURRENT FILING DATE: 2001-08-16  
PRIOR APPLICATION NUMBER: 09/666,390  
PRIOR FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 535  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-930-218-1

Query Match 37.7%; Score 1163.5; DB 9; Length 535;  
Best Local Similarity 46.2%; Pred. No. 1.1e-105;  
Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;  
QY 63 LDVSTKNPVYTNENFLSLQDPSIIHD-GWIDFLSSKRLVTIARLSAPAFIRFGKRTD 121  
DB 33 LEFYTKRPLRSVSPSLITIDASLATDPRFLTFGSLRLALRGLSPAYIRFGGTGTD 92  
QY 122 FLQFQNLKRNPAKSGPGPDYLYKNYEDDIYRSVDVALDKQKCKIAQHPDVMVLQREKA 181  
DB 93 FLIF-----DPDEPTSEERSYKSCVNHDICRSEV-----SAAVLRLQWMP 137  
QY 182 AQMHVILKEQFSNTYSNLIITARSIDKLYNPAKSGGHLIFALNLRPNNSNNSSSA 241  
DB 138 FQ-ELLILREQYQREPKNSTYRSASVDMLYSFAKSGGIDILFGNALRLTDLRNSSNA 196  
QY 242 LSLIKTSASKNTNISELGNERNYRTMGRVANSQKGYIOLKSLIOPRIYRSASL 301  
DB 197 QLLIDVCSKSGYNISWELNENPNSFWKKAHILIDQLGDEDFEVLHKLQK-SAFQNAK 255  
QY 302 YGPNIGRPKNVIALLDGFMKAGSTVDVAVTWOHCYIGRVAVKMDFLKRLDITLSQI 361  
DB 256 YGPDIGQPKGVKILRSFLKAGGEVIDSLTWHHYLNRLATKEDFLSDQDLTFIISV 315  
QY 362 RKIQVNTYTPPKKIMLEGVVTTSAGGNNLSDSYAAQFLMNTLQMLANOGIDIVIRH 421  
DB 316 QKILKVTKEITPKKWLGETSSAYGGGAPLLSNTFAAGFMWLDKLGSAQWGEIVWRO 375  
QY 422 SFPDHGYNLVDONFNP.LPDYWL.SLLYKELIGPKYLAHVAGLQRRPRGRVIRDKLRI 481  
DB 376 VFFGAGNYHLDENFRLPDYWL.SLLFKLVGPKYLAHVAGLQRRPRGRVIRDKLRI 426  
QY 482 AHCTNNHNNYVGSITFLINLHRSRKIKLAGTLRDLVHGYLQPYGQGLKSKSVQ 541  
DB 427 LHCTNYHRYGEGD.LLYVNLHNVTKILKPPPLFRKSPDVTYLLKSGPGGLSKSVQ 486  
QY 542 LMGQPLVMVDGTLPELKRPPLRAGRTLVIPVTMGFYVKNVNALAC 589  
DB 487 LMGQPLKMDQTLPLTEKPLPAGSSLSVPASISYFFVLRNKAIAAC 534

RESULT 12  
US-09-776-874A-44  
Sequence 44, Application US/09776874A  
Patent No. US20020102560A1  
GENERAL INFORMATION:  
APPLICANT: pecker, iris  
APPLICANT: vlodavsky, israel  
APPLICANT: Feinstein, Elena  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
FILE REFERENCE: 01/22603  
CURRENT APPLICATION NUMBER: US/09/776,874A  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 08/922,170

; PRIOR FILING DATE: 1997-09-02  
 ; PRIOR APPLICATION NUMBER: US 09/109,386  
 ; PRIOR FILING DATE: 1998-07-10  
 ; PRIOR APPLICATION NUMBER: PCT/US98/17954  
 ; PRIOR FILING DATE: 1998-08-31  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: Patent version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-776-874A-44

Query Match 37.7%; Score 1163.5; DB 9; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 1,1e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKMPRTVNEFLSLQDPSIIHD-GWLDPLSKRLVTLARGISPAFLRFGKRTD 121  
 DB 33 LEFTKRLRSVSPFLSITIDASLATDPRFLTGLSPRLALARGISPAFLRFGKRTD 92  
 QY 122 FLQFQNLNPAKSRGGGPPDYLLKNYEDDIYRSDVALDKQKCKIAQHPVMTLQREKA 181  
 DB 93 FLIF----DPDKPTSEERSYKSGQVNHDCRSEPV-----SAAVLRKQVEMP 137  
 QY 182 AQMHVLLKEQPSNTYSNLIITARSIDKLYNFADCSGLHIFALNALRPNNSMNSSA 241  
 DB 138 FO-ELLRLREYQKEFNKSTYSRSSVDMLYSPAKCSGLDIFGLNALRTDRLMNSNA 196  
 QY 242 LSLKYSASKKINISWEIGNEPNRYTMHGRAVNGSQKDYIOUKSLQPIRYSASL 301  
 DB 197 QLLDYCSSKGINISWEIGNEPNRSMFKKAHILIDGLQGEDFVEIHLKLR-SAFONAKL 255  
 QY 302 YGPNIGRPKRVNIALDGFPMKAGSTVDVAVTWQHCYIDGRVYKVMDFLKTLLDTSQI 361  
 DB 256 YGPDIGQPRGKTIVKLRSLRAGGEVIDSLTMHHYILNGRATYEDFLSDALDTFILSV 315  
 QY 362 RKIQKVVNTYTPGKKIMLEGVVTTSAGGTNNLSDSYAAGFLMTLGMLANQIDVYIRH 421  
 DB 316 QKILKVTKEITPGKKVMTGETSSAYGGAPLSTNTFAAGFMMIDKLGISAQMGIEVVMRQ 375  
 QY 422 SFPDHGVNHLVDQNPFLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEPFLPDYMLSLFKKLGVPRVLSRYVGPD-----RSKLRY 426  
 QY 482 AHCNTNHNHNYRGSITLFIINLHRSRKKIKLAGTLRDKLVHQYLLQPYGQGLSKSVQ 541  
 DB 427 LHCTNVHPRYQEGDLTLVNLHNTGHLKVPPLFRKPPDVTYLLKPSGPDGLSKSVQ 486  
 QY 542 LINGQPLVWVDGTLPELKRPLRAGRTLVIPVTMGFVVYKVNALAC 589  
 DB 487 LINGQILKNVDEQTLPALTEKPLPAGSALSLPAFSYGFEVIRNAKTIAC 534

RESULT 13  
 US-09-988-113-44  
 ; Sequence 44, Application US/09988113  
 ; Patent No. US20020168749A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pecker, Iris  
 ; APPLICANT: Vlodavsky, Israel  
 ; APPLICANT: Feinstein, Elena  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
 ; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS  
 ; FILE REFERENCE: 01/22781  
 ; CURRENT APPLICATION NUMBER: US/09/988,113  
 ; PRIOR FILING DATE: 2001-11-19  
 ; PRIOR APPLICATION NUMBER: US 09/776,874  
 ; PRIOR FILING DATE: 2001-02-06  
 ; PRIOR APPLICATION NUMBER: US09/258,892  
 ; PRIOR FILING DATE: 1999-03-01  
 ; PRIOR APPLICATION NUMBER: PCT/US98/17954  
 ; PRIOR FILING DATE: 1998-08-31

; PRIOR APPLICATION NUMBER: US 09/109,386  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: US 08/922,170  
 ; PRIOR FILING DATE: 1997-09-02  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: Patent version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-988-113-44

Query Match 37.7%; Score 1163.5; DB 9; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 1,1e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKMPRTVNEFLSLQDPSIIHD-GWLDPLSKRLVTLARGISPAFLRFGKRTD 121  
 DB 33 LEFTKRLRSVSPFLSITIDASLATDPRFLTGLSPRLALARGISPAFLRFGKRTD 92  
 QY 122 FLQFQNLNPAKSRGGGPPDYLLKNYEDDIYRSDVALDKQKCKIAQHPVMTLQREKA 181  
 DB 93 FLIF----DPDKPTSEERSYKSGQVNHDCRSEPV-----SAAVLRKQVEMP 137  
 QY 182 AQMHVLLKEQPSNTYSNLIITARSIDKLYNFADCSGLHIFALNALRPNNSMNSSA 241  
 DB 138 FO-ELLRLREYQKEFNKSTYSRSSVDMLYSPAKCSGLDIFGLNALRTDRLMNSNA 196  
 QY 242 LSLKYSASKKINISWEIGNEPNRYTMHGRAVNGSQKDYIOUKSLQPIRYSASL 301  
 DB 197 QLLDYCSSKGINISWEIGNEPNRSMFKKAHILIDGLQGEDFVEIHLKLR-SAFONAKL 255  
 QY 302 YGPNIGRPKRVNIALDGFPMKAGSTVDVAVTWQHCYIDGRVYKVMDFLKTLLDTSQI 361  
 DB 256 YGPDIGQPRGKTIVKLRSLRAGGEVIDSLTMHHYILNGRATYEDFLSDALDTFILSV 315  
 QY 362 RKIQKVVNTYTPGKKIMLEGVVTTSAGGTNNLSDSYAAGFLMTLGMLANQIDVYIRH 421  
 DB 316 QKILKVTKEITPGKKVMTGETSSAYGGAPLSTNTFAAGFMMIDKLGISAQMGIEVVMRQ 375  
 QY 422 SFPDHGVNHLVDQNPFLPDYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEPFLPDYMLSLFKKLGVPRVLSRYVGPD-----RSKLRY 426  
 QY 482 AHCNTNHNHNYRGSITLFIINLHRSRKKIKLAGTLRDKLVHQYLLQPYGQGLSKSVQ 541  
 DB 427 LHCTNVHPRYQEGDLTLVNLHNTGHLKVPPLFRKPPDVTYLLKPSGPDGLSKSVQ 486  
 QY 542 LINGQPLVWVDGTLPELKRPLRAGRTLVIPVTMGFVVYKVNALAC 589  
 DB 487 LINGQILKNVDEQTLPALTEKPLPAGSALSLPAFSYGFEVIRNAKTIAC 534

RESULT 14  
 US-10-341-582-44  
 ; Sequence 44, Application US/10341582  
 ; Publication No. US20030161823A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Neta Ilan  
 ; APPLICANT: Israel Vlodavsky  
 ; APPLICANT: Oron Jacoby-Zeevi  
 ; APPLICANT: It's Pecker  
 ; TITLE OF INVENTION: THERAPEUTIC AND COSMETIC USES OF HEPARANASES  
 ; FILE REFERENCE: 25449  
 ; CURRENT APPLICATION NUMBER: US/10/341,582  
 ; CURRENT FILING DATE: 2003-01-14  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: Patent version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-10-341-582-44

Query Match 37.7%; Score 1163.5; DB 14; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 1.1e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKPVATVNNENFLSLQDPSIIHD-GWLDPLSKRLVTLARGLSPAFRFGSKRTD 121  
 DB 33 LEFTYTRPLRSVSPFSITIDASLATDPRFLTFLGSPRLARGLSPAYLRFQGTXTD 92  
 QY 122 FLOFQNLRNPAKSGRGGPDYLLKNTYEDIVRSDDVALDKOKGCKIAQHPDVMVLQREKA 181  
 DB 93 FLIF-----DDKEPTSEERSYWKSQVNHDCRSEPV-----SAAVLRKLGVEWP 137  
 QY 182 AQCHLVLLKEQFSTYTNLLITARSLLDKYNPADCGSLHLIPALNLRPNNSWNSSA 241  
 DB 138 FO-ELLRLREQYQKEFNSTYSRSVDMLYSFACCSGLDLFGNALLRTPDLRMSSNA 196  
 QY 242 LSLKYSASKKYNISWEIGNEPNYRTMGRVAVNGSOLGRDYIQLKSLQPIRIYSPASL 301  
 DB 197 QLLDDYCSKYNISWEIGNEPNRSFWKKAHLLIDGLQGEDFVELHKLQR-SAFQNAKL 255  
 QY 302 YGPNIGRPRKAVIALLDGFMKAVASTVDAYTWQHCYIDGRVVKMDPLKRLDPLSDQI 361  
 DB 256 YGPDIGQPRGKTVKLLRSFLKAGGEVIDSLTWHHYINGRIATKEDPLSDALDPTFLSV 315  
 QY 362 RKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMTLNTLGMLANOGIDVYIRH 421  
 DB 316 QKILKVTKEITPGKKVWLGERTSSAYGGAPLSTPAAGFWMIDKGLSAGMGIEVVMRQ 375  
 QY 422 SFPHGYNHLYDQNFNPLPDYWSLLYKRLIGPKYLAHVHAGLQKRRPGRVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEFEPPLPDYWSLLFPKQVGPVLLSRVGPD-----RSKLRYV 426  
 QY 482 ACHTNHNNYVRGSITFLFINLHRSRKIKLAGTLDKLVHOYLLOPYQGEGLKSKSVQ 541  
 DB 427 LHCTNVYHPRYQEGDITLYVNLHNVTGHLKVPPLPRKPDYTYLLKPSGPDGLSKSVQ 486  
 QY 542 LINGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVKNVNALAC 589  
 DB 487 LINGQILKMWDEQTLPALTEKPLPAGSALSLPASFYGFVJNNAKIAAC 534

RESULT 15  
 US-10-384-451-44  
 ; Sequence 44, Application US/10384451  
 ; Publication No. US20030170860A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pecker, Iris  
 ; APPLICANT: Vlodavsky, Israel  
 ; APPLICANT: Feinstein, Elena  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
 ; FILE REFERENCE: 25718  
 ; CURRENT APPLICATION NUMBER: US/10/384,451  
 ; CURRENT FILING DATE: 2003-03-10  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 44  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-384-451-44

Query Match 37.7%; Score 1163.5; DB 14; Length 535;  
 Best Local Similarity 46.2%; Pred. No. 1.1e-105;  
 Matches 244; Conservative 80; Mismatches 177; Indels 27; Gaps 6;

QY 63 LDVSTKPVATVNNENFLSLQDPSIIHD-GWLDPLSKRLVTLARGLSPAFRFGSKRTD 121  
 DB 33 LEFTYTRPLRSVSPFSITIDASLATDPRFLTFLGSPRLARGLSPAYLRFQGTXTD 92  
 QY 122 FLOFQNLRNPAKSGRGGPDYLLKNTYEDIVRSDDVALDKOKGCKIAQHPDVMVLQREKA 181

DB 93 FLIF-----DDKEPTSEERSYWKSQVNHDCRSEPV-----SAAVLRKLGVEWP 137  
 QY 182 AQCHLVLLKEQFSTYTNLLITARSLLDKYNPADCGSLHLIPALNLRPNNSWNSSA 241  
 DB 138 FO-ELLRLREQYQKEFNSTYSRSVDMLYSFACCSGLDLFGNALLRTPDLRMSSNA 196  
 QY 242 LSLKYSASKKYNISWEIGNEPNYRTMGRVAVNGSOLGRDYIQLKSLQPIRIYSPASL 301  
 DB 197 QLLDDYCSKYNISWEIGNEPNRSFWKKAHLLIDGLQGEDFVELHKLQR-SAFQNAKL 255  
 QY 302 YGPNIGRPRKAVIALLDGFMKAVASTVDAYTWQHCYIDGRVVKMDPLKRLDPLSDQI 361  
 DB 256 YGPDIGQPRGKTVKLLRSFLKAGGEVIDSLTWHHYINGRIATKEDPLSDALDPTFLSV 315  
 QY 362 RKIQKVNNTYTPGKKIMLEGVVTTSAGTNNLSDSYAAGFLMTLNTLGMLANOGIDVYIRH 421  
 DB 316 QKILKVTKEITPGKKVWLGERTSSAYGGAPLSTPAAGFWMIDKGLSAGMGIEVVMRQ 375  
 QY 422 SFPHGYNHLYDQNFNPLPDYWSLLYKRLIGPKYLAHVHAGLQKRRPGRVIRDKLRIY 481  
 DB 376 VFFGAGNYHLVDENEFEPPLPDYWSLLFPKQVGPVLLSRVGPD-----RSKLRYV 426  
 QY 482 ACHTNHNNYVRGSITFLFINLHRSRKIKLAGTLDKLVHOYLLOPYQGEGLKSKSVQ 541  
 DB 427 LHCTNVYHPRYQEGDITLYVNLHNVTGHLKVPPLPRKPDYTYLLKPSGPDGLSKSVQ 486  
 QY 542 LINGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVKNVNALAC 589  
 DB 487 LINGQILKMWDEQTLPALTEKPLPAGSALSLPASFYGFVJNNAKIAAC 534

Search completed: May 6, 2004, 13:51:38  
 Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

## OM protein - protein search, using sw model

Run on: May 6, 2004, 13:42:11 ; Search time 20 seconds  
(without alignments)  
2847.269 Million cell updates/sec

Title: US-10-088-676-2

Sequence: 1 MRTLCAFPPEAMPSSNSRPPA.....PYTMGFYVKNVNALACRYR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: PIR.78:\*  
2: PIR.1:\*  
3: PIR.3:\*  
4: PIR.4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2446	79.2	480	2 JCT506	heparanase protein
2	382	12.4	521	2 T45608	hypothetical prote
3	155	5.0	190	2 T01953	hypothetical prote
4	146	4.7	1260	2 T14022	reverse transcript
5	128	4.1	1027	2 AC1841	glycerophosphoryl
6	123	4.0	2026	1 OYBY	adenylate cyclase
7	115	3.7	923	2 S50482	hypothetical prote
8	114.5	3.7	837	1 A31842	endo-1,4-beta-xyla
9	106.5	3.4	1839	1 OIBIK	adenylate cyclase
10	106	3.4	484	2 C88264	protein kin-15 (lm
11	106	3.4	484	2 T14430	hypothetical prote
12	105	3.4	747	2 S66959	hypothetical prote
13	105	3.4	835	2 C97322	probable alpha-ara
14	104.5	3.4	839	2 S73548	MG422 homolog C12
15	104.5	3.4	1280	2 B34087	hypothetical prote
16	104	3.4	6350	2 T31679	heparanase synthet
17	101	3.3	535	2 S58740	cytochrome-c oxida
18	101	3.3	779	2 B42375	endopeptidase la
19	101	3.3	2244	2 F90563	hypothetical prote
20	101	3.3	2971	2 T08026	hypothetical prote
21	100.5	3.3	498	2 F83329	probable outer mem
22	100.5	3.3	566	2 T13042	hypothetical prote
23	100.5	3.3	630	2 C71374	probable glucose i
24	100.5	3.3	878	2 A55201	meiosis-specific p
25	100.5	3.3	1259	4 GNMU1	retrovirus-related
26	100.5	3.3	3343	4 S44887	ZK112.7 protein -
27	99	3.2	785	2 F97001	endoglucanase fami
28	99	3.2	785	2 B72608	probable hyua APE1
29	99	3.2	871	2 S68482	probable membrane

30	98.5	3.2	1356	2 S51389	ROM2 protein - yea
31	98	3.2	629	2 C64180	hypothetical prote
32	98	3.2	830	2 C86191	hypothetical prote
33	98	3.2	866	1 NCBCX8	exodeoxyribonuclea
34	98	3.2	866	2 F90870	exonuclease VIII R
35	98	3.2	866	2 C85748	exonuclease VIII R
36	98	3.2	912	2 T18785	hypothetical prote
37	98	3.2	1047	2 C87358	AcirB/AcirD/AcirF
38	97.5	3.2	630	2 C84453	probable selenium-
39	97.5	3.2	1607	2 T04583	TMV resistance pro
40	97	3.1	385	1 T52057	trehalase-phosphat
41	97	3.1	543	2 T37570	WD repeat protein
42	97	3.1	897	2 F71816	DNA polymerase I -
43	97	3.1	1391	2 T20642	hypothetical prote
44	97	3.1	1397	2 S87998	protein F09C3.1 [1
45	97	3.1	1661	2 T21986	hypothetical prote

## ALIGNMENTS

## RESULT 1

JCT506

heparanase protein 2a - human

C/Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 01-Dec-2000

C/Accession: JCT506

R/McKenzie, E.; Tyson, K.; Stamps, A.; Smith, P.; Turner, P.; Barry, R.; Hircock, M.; Pat

Biochem. Biophys. Res. Commun. 276, 1170-1177, 2000

A/Title: Cloning and expression profiling of Hpa2, a novel mammalian heparanase family m

A/Reference number: JCT506

A/Accession: JCT506

A/Molecule type: mRNA

A/Residues: 1-480 &lt;MCK&gt;

A/Cross-references: GB:AF282885

C/Comment: This protein, a intracellular membrane-bound enzyme, has biological and therai

therapies.

C/Genetics:

A/Gene: hpa2a

A/Map position: 10q23-10q24

C/Keywords: heparin binding; membrane bound

Query Match 79.2%; Score 2446; DB 2; Length 480;

Best Local Similarity 80.9%; Pred. No. 1.5e-183;

Matches 479; Conservative 1; Mismatches 0; Indels 112; Gaps 1;

QY	1	MVTLCAFPPEAMPSSNSRPPACIAPALYLAIIHISSSQAGDRRLPVDRAGKXETL	60
DB	1	MVTLCAFPPEAMPSSNSRPPACIAPALYLAIIHISSSQAGDRRLPVDRAGKXETL	60
QY	61	ILDDVSTKNPVRTVVENFLSLQDPSIHDGWLDFSSKRLVTLARGSPAFLEFGKRT	120
DB	61	ILDDVSTKNPVRTVVENFLSLQDPSIHDGWLDFSSKRLVTLARGSPAFLEFGKRT	120
QY	121	DLQFQNTNRPKASGSGGPGDYLLKNYEDDIVRSVVALDKQCKIAQHPDMLVLOREK	180
DB	121	DLQFQNTNRPKASGSGGPGDYLLKNYEDDIVRSVVALDKQCKIAQHPDMLVLOREK	180
QY	181	AAQWMLVLKQGFSTVYSNLIILARSCLKVNFADCSGLHIFALNALRPNNSWNSSS	240
DB	181	AAQWMLVLKQGFSTVYSNLIILARSCLKVNFADCSGLHIFALNALRPNNSWNSSS	240
QY	241	ALSLKYSASKYNTISWEIGNEPNRYTMHGRAVNGSLQKDYQLKSLQPIRIRYRAS	300
DB	241	ALSLKYSASKYNTISWEIGNEPNRYTMHGRAVNGSLQKDYQLKSLQPIRIRYRAS	300
QY	301	LVGPNIGRPKNVIALLDGEMKVAAGSTDAVYWKQCYIDGRVVKMDLKTRLDPTLSDQ	360
DB	301	LVGPNIGRPKNVIALLDGEMKVAAGSTDAVYWKQCYIDGRVVKMDLKTRLDPTLSDQ	360
QY	361	IRIKQKVVNTYTPGKKIMLEGVVITSAGGTNNLSDSYAAGFLMNTLGLANOGIDVIR	420
DB	361	IRIKQKVVNTYTPGKKIMLEGVVITSAGGTNNLSDSYAAGFLMNTLGLANOGIDVIR	420
QY	420	IRIKQKVVNTYTPGKKIMLEGVVITSAGGTNNLSDSYAAGFLMNTLGLANOGIDVIR	480
DB	420	IRIKQKVVNTYTPGKKIMLEGVVITSAGGTNNLSDSYAAGFLMNTLGLANOGIDVIR	480

```

QY 421 HSPFDHGNHLYVDQNFPELPDYMLSLKYRLIGPKYLAHVAGLQKRPGRVIRDKRL 480
DB 309 HSPFDHGNHLYVDQNFPELPDYMLSLKYRLIGPKYLAHVAGLQKRPGRVIRDKRL 368
QY 481 YAHCTNHNHNHYVGSITLFTINLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLSKSV 540
DB 369 YAHCTNHNHNHYVGSITLFTINLHRSRKKIKLAGTLRDLVHQYLLQPYGQGLSKSV 428
QY 541 QLNQGPLVWDDGTLPELKPRLPAGRTVIRPTWGFVYKVNVLACGR 592
DB 429 QLNQGPLVWDDGTLPELKPRLPAGRTVIRPTWGFVYKVNVLACGR 480

RESULT 2
T45608
hypothetical protein F13G24.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45608
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voel, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, December 1999
A:Reference number: 223009
A:Accession: T45608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-521 <BEV>
A:Cross-references: EMBL:AL133421
A:Experimental source: cultivar Columbia; BAC clone F13G24
C:Genetics:
A:Map position: 5
A:introns: 53/3; 66/1; 127/2; 177/1; 256/1; 319/2; 361/2; 394/3
A>Note: F13G24.30

Query Match 12.4%; Score 382; DB 2; Length 521;
Best Local Similarity 25.8%; Pred. No. 8.1e-22;
Matches 154; Conservative 76; Mismatches 187; Indels 180; Gaps 28;

QY 75 NENLSLQLD--PEIHD-----GWLDFLS--SKRLVTLA-RGLSPALRGGGR 119
DB 22 DENFVCATLDMWP--HDKNYDQCPWGSVIMDLTRPLTLTAIAPFLRIRIGSL 78
QY 120 TDFLOFO--NLKNDP--KSRGPGPDYLLKNYEDDIVSDVALDKQCKIAQHPDV 173
DB 79 QDQYIVDVGNLKTCPRFQKNGS-----LFGSKXC----- 110
QY 174 LVLRERAAQMHVLLKQPSNTYSNLLTRASLDKLYNADSGHLITALNLR 229
DB 111 -----LHM-----KRWDELNSFLTATGAVVTGALNLRGRHK 142
QY 230 ---RNPNNNSNSALSLKYSASKKTYNI--SWEIGNEPNNYRTWNGRAVNGSOLGKDYI 285
DB 143 LRGKAWGAMWHINTOPFLANTYVSKGYVIDSMERGNLSGSV--GASVABELXGDLIV 200
QY 286 LKSLQIPRIYRSALSYGPNIGRPRKRVIALDGF-----MKVAG--STDVAVTWQR 335
DB 201 LKQVIN--KYKNSWMLKHPILVAP-----GGFYEQQWYTKLEISGSDVDVAV--RH 249
QY 336 CYIDGR-----VYKWMDFLKTLLDLSDOIIRKIQKVVNTYPPGKKIMLEGVTTISAG- 388
DB 250 IYNGSGNDPLVLYKIND--PSYLSQVSKTFKQVNTQIENGP-----MASPWGSGGA 302
QY 389 ---GTNNLSDSYAAGFLMLNTLGMLANOGDIVIRHSPFDHGNHLYVDQNFPLDYMLS 445
DB 303 YNSGRHVSDFTFISFWYLDQIGMSARHNTKYVCRQTLVGGFYGLKKGFFVNPDPDYSA 362
QY 446 LLVYRLIGPKYLAHVAGLQKRPGRVIRDKRLIYAHCTNHNHNHYVGSITLFTINLH 505
DB 363 LLVHRLMGKYLAVQTGD-----PQLAVYAHCSG-----PAGVTILLNLNS 405
QY 506 -----RSRRKKIKLAGTLRDLV-----HQYLLQPYGQ 533
DB 406 NQSDFTVSVSNGINVLNABSRKKSLDLTLKRPFSWIGSKASDGYINREYHLLTP--EN 463

```

```

QY 534 G-LKSKSVQLNGQPLVWDDGTLPELKPRLPAGRTVIRPTWGFVYKVNVLAC 589
DB 464 GVLRSKTMVNLGSKLKPATNDISLEPVRSVNSPLNVLPLSSFLVLRNPDASAC 520

RESULT 3
T01953
hypothetical protein T2L5.6 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jan-2000
C:Accession: T01953
R:Geisel, C.; Smith, A.; Le, T.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of A. thaliana T2L5.
A:Reference number: 214470
A:Accession: T01953
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-190 <GEI>
A:Cross-references: EMBL:AF096371; NID:g3695386; PID:g3695392
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:introns: 36/2; 69/3
A>Note: T2L5.6
C:Superfamily: Arabidopsis thaliana hypothetical protein T2L5.6

Query Match 5.0%; Score 155; DB 2; Length 190;
Best Local Similarity 24.9%; Pred. No. 0.0001;
Matches 49; Conservative 32; Mismatches 70; Indels 46; Gaps 6;

QY 420 HSPFDHGNHLYVDQNFPELPDYMLSLKYRLIGPKYLAHVAGLQKRPGRVIRDKRL 479
DB 12 RQSLIGNVGLNTNTPFPNDPYSALTWRLQMRKALFTTFSGTK-----KIR 60
QY 480 IYAHCTNHNHNHYVGSITLFTINLHRSR--KKIKLAGTLRDLVHQYLLQPY----- 530
DB 61 SYTHCARQSK-----GITVLMLNDNTTVAVAVELNNSP--SLRHTKHKYSKRASQ 112
QY 531 ---GQEG-----LKSQVQLNGQPLVWDDGTLPELKPRLPAGRTVIR 572
DB 113 LFGGPNGYIQREHYLTAKDGNLHSDTLNGLNQNWSGDLPIEIRHINSPTITIA 172
QY 573 PYTMGFYVYKVNVLAC 589
DB 173 PYSLVFYHMRNVVVPAC 189

RESULT 4
T14022
reverse transcriptase homolog - slime mold (Dictyostelium discoideum) retrotransposable
C:Species: Dictyostelium discoideum
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T14022
R:Winkler, T.; Tschepke, C.; de Hostos, E.L.; Jendretzke, A.; Dingermann, T.
Mol. Gen. Genet. 257, 655-661, 1998
A:Title: Tdd-3, a transfer RNA gene-associated poly(A) retrotransposon from Dictyostelium
A:Reference number: 217858; MUID:98265925; PMID:9604889
A:Accession: T14022
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1260 <MIN>
A:Cross-references: EMBL:AF002669; NID:g2558820; PID:g2558822; PIDN:AA048324.1

Query Match 4.7%; Score 146; DB 2; Length 1260;
Best Local Similarity 21.3%; Pred. No. 0.0099;
Matches 141; Conservative 101; Mismatches 229; Indels 192; Gaps 39;

QY 63 NPVRTVNEFLSLQDPSIHDGWL---DFL---SKRLV-----TLARGSP 110
DB 94 NGIGLHNHNNQIKSLPFIIEGRLLISDLIKOTTRTLAIYAPQPKKRTIASLTK 153

```



```

QY 111 AFLFGGKRTDPLQFQNTLRNPAKSGGPGPDYLYKNVEDDIYRSVVALDKOKCKLIAOH 169
DB 154 HFNK-----QYHULTSNPKNKNIDIIAGDFNCIDPNDNHTSND-----DQGNLTQS 199
QY 170 PDVVLVYOR-----EKAQWHLVLLKEQFS-----NTYSNLI-LTARSLDKLY-----NF 213
DB 200 PDEMAVTEAIRISNNIMDMQJLNKRPTFSRTINTNNLTTRILERRLDRIYANNSILNY 259
QY 214 -----ADCSGLHIFALNALARNNNMW--NSSGALS--LLKXSASK 251
DB 260 SOLYRNLIIPKINDIPLSDFNFLTFTLANIQTNRR--WRLKSSILSSIMLK----- 313
QY 252 KYNISWELGNEPNRYRTMGRVAVNGSQIGKDYIQLKSLQPIR-IYSRASLYGPNIGRPR 310
DB 314 --NIDFL-----NGY-----SRELSNHNISFSQJNSLKNKIKQLYTEFQKQDYNKAN 363
QY 311 --KAVIALD--GFMVAVASTYDAYTW-----QHC-----YIDGRVVK--VMDP 348
DB 364 IKNLISLETFEYKQAFATLSAINESKREBQJLKOELANNCEETSLKYISARIKKRNDP 423
QY 349 LKTRLDLTLSQJRIQKVNTYT-----PGKTIWLGAVTTSAGTNNLS 394
DB 424 TINAVKOTQGTINKQELIEEYKYXSNLYDYKEDDPESHYELEWVYTRDSTWMDLE 483
QY 395 DSYAGFLW--LNTGLMLANOGIDVTRHSEFDD-----GYN-----HLYDQNF 436
DB 484 NEFTQSEILEVYKQJLNPKHSGPDPGIPMLFITHKEKLAFLASAFNDLTRNPHLISKNY 543
QY 437 -----NPLPDYMLSLVK-----RLIGPKYLAHVHAGLQKRR-----PGR 412
DB 544 KEGIIITPKKQDEPLKRRPITLANCIYKISHKLNNEIIPITLYVINHNOXGPFPR 603
QY 473 VIRDKLRIYACHTNHNHNVYSGITLFIINLHRSRKKIKLAGTRDKL--VHOYLLQPY 530
DB 604 FILNITISINELIYXCDKRIINGIITLYI-----SKLLTRSHVQSQITTHQHSQYI 658
QY 531 GQ-BGLSKS---VOLNGQPLVWDDGTLR-ELKPRLRAGRTVIRPYVMGYVYKXNV 555
DB 659 NLINMLTKSEARLEINGRT-----TIPEIK-RGVKQG-----DPLSPFLFLV-IE 704

QY 586 ALA 588
DB 705 ALA 707

RESULT 5
AC1841
glycerophosphoryl diester phosphodiesterase [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AC1841
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759640
A/Accession: AC1841
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1027 <CUR>
A/Cross-references: GB:BA000019; PIDN:BA077799.1; PID:G17135253; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: all0275

Query Match 4.1%; Score 128; DB 2; Length 1027;
Best Local Similarity 21.9%; Pred. No. 0.18;
Matches 127; Conservative 70; Mismatches 222; Indels 160; Gaps 30;

QY 62 LLDVSTKNPRTVNE-----NFLSQDPSII-HDGMIDFLSKRLVTIAGLSAPFL 113
DB 101 LSDPNNKIPQIYVNEASADRLITGADLDISFVIDXG-----SIWVGDFEFGYLL 151

```

```

QY 114 RFG--GKRTDFL-----OFQNLRNPAKS-----RGSPG--PDYLYKNVE-----DIYRS 154
DB 154 HFDATGKLIDAPLAPPRFKTLIDGTAPREVIGHRASGRPREHITLESYLAIEGADFIER 211
QY 155 DVALDKOKGCKIAQHPDVLVYOREKAQWHLVLLKEQFSNTYSNLI-LTARSLDKLYN-- 212
DB 212 DLAVTKD--GVILARHEPALAVLNDGVS-----NFSNTTNVQIAPFSRLKTVN 261
QY 213 -----PADCSGLHIFALNALARNP--NNSNASSALSILKYSASKKNYSWELGN 261
DB 262 LDGTEITGFAEDFTLAEIKELRAIERLFRDQSGNPGFTTAEIIDLKQVAEAFKQ 321
QY 262 EPNRY-RTWGRVAV-----NSQJCKDYIQLKSLQPIRISRASLYGPN 305
DB 322 KIGIYPERKHPYFQAGATVYGTTEKINRNISQIILDKANNFTDPERITIGSEVQ-- 379
QY 306 IGRPRKNVIALIDGEMKAVAGSTV-----DAVTWQCYIDGRVVK--MDFLKTRLDLT 357
DB 380 -----NLKEHDTIMPAGVDIPLVQLFDAIDVD--INGRLIETREYDFIVSGDRTTY 430
QY 356 SDQIRKIQV--VNTYTPGKIMLEGVVTSAGTNNLSDSYAGFLMLNTGLMLANOGI 415
DB 431 GD-LTPAGLAIATYADGIGPKRMIVSAGTDANN-----DQAD 471
QY 416 DVVIRHSEFDDGYNVLVDQNFPLPDYMLSLYKRLIGPKYLA--VHYAGLQKRRPGRV 473
DB 472 DV-----NGDGAVNDAD-----KTLLEPFLTVQDAHVHAGVQVHP---YT 507
QY 474 IRDKLRIYACHTNHNH--NY-----VSGITLFIINLHRSRKKIKLAGT--LR--- 518
DB 508 FRDEERYLA--ANYQNGELEFYQQLGCVDALFDFEPTIADRRDRSLSPGNIVRSPQ 565
QY 519 --DKLVHGYLLQPYGOEGLSKSVOLNGQPLVWDDGTLR 555
DB 566 NPDVLSGDAFNLGGSRFEGGAINASKTKLYMLEGTV 604

RESULT 6
OXYB
adenylate cyclase (EC 4.6.1.1) - yeast (Saccharomyces cerevisiae)
N/Alternate names: ATP pyrophosphate-lyase; protein J1401; protein YJL005W
C/Species: Saccharomyces cerevisiae
C/Date: 28-Dec-1987 #sequence_revision 08-Sep-1995 #text_change 21-Jan-2000
C/Accession: S56776; S56775; A24776; S05828; S55183
R/To Van, D.; Pereira, J.; Jacq, C.
submitted to the Protein Sequence Database, September 1995
A/Reference number: S56776
A/Accession: S56776
A/Molecule type: DNA
A/Residues: 1-1823 <DEH>
A/Cross-references: EMBL:Z49280; GSPDB:GN00010; MIPS:YJL005W
R/Kataoka, T.; Broek, D.; Wiggler, M.
Cell 43, 493-505, 1985
A/Title: DNA sequence and characterization of the S. cerevisiae gene encoding adenylate
submitted to the Protein Sequence Database, September 1995
A/Reference number: S56775
A/Accession: S56775
A/Molecule type: DNA
A/Residues: 673-2026 <ZAG>
A/Cross-references: EMBL:Z49280; GSPDB:GN00010; MIPS:YJL005W
R/Kataoka, T.; Broek, D.; Wiggler, M.
Cell 43, 493-505, 1985
A/Title: DNA sequence and characterization of the S. cerevisiae gene encoding adenylate
submitted to the Protein Sequence Database, September 1995
A/Reference number: A24776; MUID:86079531; PMID:2934138
A/Accession: A24776
A/Molecule type: DNA
A/Residues: 1261,'U', 263-547,'U', 549-591,'H', 593-708,'I', 710-961,'P', 963-1387,'S', 1389-
A/Cross-references: EMBL:M12057; NID:G171359; PIDN:AA4549.1; PID:G171360
A/Note: the authors translated the codon TTA for residue 262 as Ser, AGC for residue 311
2 as Leu, TCA for residue 1566 as Ala, GAG for residue 1659 as Ala, GGT for residue 1735
R/Asson, P.; Lenzen, G.; Jacquemin, J.M.; Danchin, A.
Curr. Genet. 10, 343-352, 1986
A/Title: Yeast adenylate cyclase catalytic domain is carboxy terminal.
A/Reference number: S05828; MUID:88165073; PMID:3327602

```

A/Accession: S05828  
 A/Molecule type: DNA  
 A/Residues: 1042-1426, 'D', 1428-1460, 'T', 1462-1955, 'V', 1957-2008, 'W', 2014-2015, 'TNYKMLRT'  
 A/Cross-references: EMBL:X03449; NID:G3449; PIDN:CAA2175.1; PID:G3488  
 A/Date Haan, M., Smits, P.H.M., Grievell, L.A.  
 submitted to the EMBL Data Library, May 1995  
 A/Reference number: S55183  
 A/Accession: S55183  
 A/Molecule type: DNA  
 A/Residues: 673-2026 <DEM>  
 A/Cross-references: EMBL:X87611; NID:G854567; PIDN:CAA60917.1; PID:G854568  
 A/Genetics:  
 A/Genes: SGD:CYR1; CDC35; MIPS:YJL005W  
 A/Cross-references: SGD:S0003542; MIPS:YJL005W  
 A/Map position: 10L  
 C/Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolog  
 C/Keywords: CAMP biosynthesis; duplication; phosphorus-oxygen lyase; tandem repeat  
 F/669-1343/Region: leucine-rich 23-residue repeats  
 F/1065-1087/Domain: leucine-rich alpha-2-glycoprotein repeat homolog <LRR>  
 F/1610-2026/Domain: catalytic #status predicted <CAT>  
 F/1664-1743/Domain: yeast adenylate cyclase catalytic domain homology <YACC>

Query Match 4.0%, Score 123, DB 1, Length 2026;  
 Best Local Similarity 22.6%, Pred. No. 1.3, Index 162, Gaps 33,  
 Matches 141, Conservative 83, Mismatches 237,  
 37 LSSQAGRRPLPVDRAAGLKE-KTLLLDVSTKPVG--RTVNEFLSLQDPSIIHDGW 92  
 891 LNLQCNLESLP---AGFVELKNLQDLDSNNKFMHPEVINYCNLQIDLSYKIKG 946  
 93 LDFLSSRLVTLAE-GISPAFLRFGKRTDF--LQFQNR-----NPAKSRGGP 140  
 947 LP-OSTYLYLAKMNSHKLNFIIDSEMTDLRTLNLRNRSIKTNASVLQMLFLT 1005  
 141 DYYLYKVEDIVRSADVLDKQCKIAQHPDVLVLOREKAQM-HLVLLKEQFSYTSN 139  
 1006 DNRISNED-----TLPLKRALEIOENPTISFDFPKMTSLTLNKAQLSSIPGE 1058  
 200 LILTARSLDKLYNPAACSGHLIFALNALRPNPNNSNSSLKLYSASKYNTSMEL 259  
 1059 LITLTSLEKLE-----LNQNNLTFLPOEI-SKLTQVFLSVANNKLEYIPPEL 1106  
 260 GNEENNYRT--MGRAYNGSLQKDYQLKSLQPIRTYRASL-----YGPNIQR 308  
 1107 -SUKSLFTLDLSNNIRDFPDGMENTELISLINSNAFGNSSLENSFYHMYGSLGSK 1165  
 309 PRKNVIALDGFPMKAVASTVDAVTWQ--HCYIDGRVVKVDFLTKRLDITLSDQIRIK 366  
 1166 -----SLM--FLIADNQPDAMWPLNCFVN---LKVNLSTYNNFSDV---SHMKLES 1211  
 367 VVNTYTPQKIKWLEGVTTTSAGTNNLSDSYAAQFLW--LNTLGMLANQIDV-----V 418  
 1212 ITTELISGNKL-----TTLSGDT-----VLKNSLTKLMLNSNMJSLPAELSNL 1256  
 419 IRHSFFDHVNLVDONENPLPDY--WLS--LTYKRLIGP----- 454  
 1257 SOLSPFDVGNALCKNTSNHYDMWNRNKKELKTLNFSGNRFEIKFSISHDIDADLSL 1316  
 455 -----KTLAVHVAQKQKRPGRVIRDKLRIYAHCTNNHNN-----NTY----- 493  
 1317 TVLPQKLKVLGMDVTLNTTKVDPDENVFRLETTAIIIGMKRYGADTLGQRDYSSSDVT 1376  
 494 -----RGSITLFIINLHSRRK-----IKLAGTLRK-LVHQYLLQPYGOE----- 533  
 1377 FERRRGNDDECLCHHSKNQADYGHNISRIYVDYDKILIRQ--LERVGDEDDNIXT 1434  
 534 GLKSKSVOLNGQ---PLVMVDG 553  
 1435 ALRPSFLQNLKINGMANSVDNG 1457

RESULT 7  
 S50482

Hypothetical protein YER024W - Yeast (Saccharomyces cerevisiae)  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002  
 C/Accession: S50482  
 A/Accession: S50482  
 A/Molecule type: DNA  
 A/Residues: 1-923 <DIR>  
 A/Cross-references: EMBL:U18778; NID:G603592; PID:G603616; GSPDB:GND0005; MIPS:YER024W  
 A/Genetics:  
 A/Genes: SGD:YAT2; MIPS:YER024W  
 A/Cross-references: SGD:S0000826  
 A/Map position: 5R

Query Match 3.7%, Score 115, DB 2, Length 923;  
 Best Local Similarity 18.7%, Pred. No. 1.6, Index 206, Gaps 27,  
 Matches 115, Conservative 79, Mismatches 216,  
 33 LHSQAGRRPLPVDRAAGLKEKTLLLDVSTKPVTVNEFLS--LQDPS--II 88  
 98 LHDINNQSTR-----IQDDVLRPPLVLADALPNTIQADRSALV 142  
 89 HDGWLDFLSSRLVTLARG--LSPALRFGKRTDFLOFQNL---RNPAKSRGG----- 137  
 143 H-----SAARFISLQDILLPPDINATNGKPLSMAPFLNLFQTTSPVFORGEVENFD 195  
 138 -----GQPDV-----YLKRYEDDIVRSADVLDKQCKIAQHPDVLVLO 177  
 196 LNKPTASDLDLDPYSSDNDNDPEQDPPDRKRKHEDFTFGNGITIKHPSPSKHILI 255  
 178 REKAQOMELVILKEQFSNTYNSLITPARSIDLKYN--PACSGHLIFALNAL----- 228  
 256 ISRGQYLLRYLV-----DSTWKIITYTAEITTFIRNHIKSSGGEKSTALGSLTSHSFRN 310  
 229 -----RPNPNNSNSSLKLYSASKYNTSMELNEPNRYTHMGRAYNGSL 279  
 311 WKYARKLQKRYVLEHRLIDSLFVLVDEQERTN--DSDDTADISQMNRRITTRDK 368  
 280 GROYLQKSLQPIRYERASLYGENTIGRPKNVIALDGFPMKAVASTVDA--VTWQHC 336  
 369 KCTSANCK-----RVFGTGISNKGHVGSCSRWYDKQLVW--TADAKATVINDSF 420  
 337 YIDGRVVKVDFLTKRLDITLSDQIRIKQKVVNTYTPGKIY-----LEGVV 383  
 421 TCDGSV--VLRFTS---EITYESVRLARDVNAQDPQFSIMPNVTQMDPEYTKLMATI 474  
 384 TTSAGTNNLSDSYAAQFLWNTLGMLANQIDVIVIRHSFFDHVNLVDONENPLPDYW 443  
 475 SADGGSPSIDPK-----LVVNK-----IDWSFS----- 498  
 444 LSLIKRLIGPVLAHVHVAQKQKRPGRVIRDKLRIYAHCTNNHNNHYVASTLFIIN 503  
 499 -----NINLTHVHLETK-----LADLSKI-----DIYASITPL----- 528  
 504 LHSRKKIKLAGTLRDLKVHGYL-----LQPYGQBLK-SKSVOLNGQPLV 548  
 529 -----GRRSAGRLGVKDSNVQVALQIAHYALYGMVFGLEPSTGFKNSRSSFINIQQA 585  
 549 MYD-----DGT 554  
 586 LTELQQLTSSSIDGT 601

RESULT 8  
 A31842  
 endo-1,4-beta-xylosidase (EC 3.2.1.8) Z precursor - Clostridium thermocellum  
 N/Alternate names: xylosidase Z  
 C/Species: Clostridium thermocellum  
 C/Date: 31-Mar-1990 #sequence\_revision 11-Apr-1997 #text\_change 18-Jun-1999  
 C/Accession: A31842

R,Greiner, O.; Chebrou, M.C.; Beguin, P.  
 J. Bacteriol. 170, 4582-4588, 1988  
 A>Title: Nucleotide sequence and deletion analysis of the xylanase gene (xynZ) of Clostridium  
 A/Reference number: A31842; MUID:8908072; PMID:3139632  
 A/Accession: A31842  
 A/Molecule type: DNA  
 A/Residues: 1-837 <GRE>  
 A/Cross-references: GB:M22624; NID:g144931; PIDN:AAA3286.1; PID:g144932  
 C/Genetics:  
 A/Status: xynZ  
 A/Function:  
 A>Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans  
 C/Pathway: xylan degradation  
 C/Superfamily: Clostridium endo-1,4-beta-xylanase Z; Clostridium cellulase repeat homolo  
 C/Keywords: duplication; extracellular protein; glycosidase; heat-stable protein; hydrol  
 F/1-28/Domain: signal sequence #status predicted <SIG>  
 F/29-837/Product: endo-1,4-beta-xylanase #status predicted <MAT>  
 F/326-419/Domain: Clostridium xylanase A repeat homology <CXA>  
 F/430-453/Domain: Clostridium cellulase A repeat homology <CCR1>  
 F/464-487/Domain: Clostridium cellulase repeat homology <CCR2>  
 F/548-834/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>  
 F/645,754/Active site: Glu #status predicted

Query Match 3.7%; Score 114.5; DB 1; Length 837;  
 Best Local Similarity 21.0%; Pred. No. 1.5;  
 Matches 77; Conservative 54; Mismatches 128; Indels 107; Gaps 20;

QY 128 LBNPARKS---GGPPDYIKNYEDDIVRSDVALDKQCKIAQHPDVMVLQREKAAQM 184  
 DB 518 LRDYAEARKGKIGTCVNYPPYNSDPYNS-----LQREPSMVCNE----- 561  
 QY 185 HLVLKEQPSNTYSNLLITARSIDKLVNFDSCGL---HLIFALNALRNP---NSM 236  
 DB 562 ---MKFDALQPRQNVDPFSGK--DQLLAFARNGMQRGHTLIWHN---QNPFWLITNGW 613  
 QY 237 NSSALSILKYSAS-----KKNISWELGNE-----PNNYRMHGRVAVSGLGKDIQ 285  
 DB 614 NRSLAIVKQNHITTVTHYKGIKIVDVANECMDSGNLSIMKNV---IGDDYLD 669  
 QY 286 LKSLQPRIRIYSPAS---LYG---PNIGRPKNVIALDGFVKAGSTVDVATWQH 335  
 DB 670 YAPF---YARADADALLFNVDVIEDLG--PKSAVAVNMKSKMERGVPIIDGVGFC 722  
 QY 336 CYIDGRVVKMDFKITLTLTSDQIRKIQV-----VNTYTPGKTIWLEGVYTTSA 387  
 DB 723 HFIING-----MSPEYLASIDONIKRYAIGVYSFTEIDIRIPQS---ENPATAFQ 770  
 QY 388 GGTNNLSDSYAAQFLMNTIGMLANGIDIVIRHSFDDH-----GYNH--LVDPNF 436  
 DB 771 VQANNYELMKI-----CLANPNONTFVWNGFTDKTWIPIGTPPGYGNFLYDSNY 821  
 QY 437 NPLPDY 442  
 DB 822 NPKPAY 827

## RESULT 9

OYBYK  
 adenyLate cyclase (EC 4.6.1.1) - yeast (Saccharomyces kluyveri)  
 N/Alternate names: adenylyl cyclase  
 C/Species: Saccharomyces kluyveri  
 C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 21-Jan-2000  
 C/Accession: J01145; S14464  
 R/Young: D.; O'Neill, K.; Broek, D.; Wiegler, M.  
 A>Title: The adenylyl cyclase-encoding gene from Saccharomyces kluyveri.  
 A/Reference number: J01145; MUID:91323718; PMID:1864503  
 A/Accession: J01145  
 A/Molecule type: DNA  
 A/Residues: 1-1839 <YOU>  
 A/Cross-references: EMBL:X56042; NID:g4856; PIDN:CAA39513.1; PID:g4857  
 C/Genetics:  
 A/Status: CYR1  
 A/Gene: CYR1

C/Superfamily: yeast adenyLate cyclase; leucine-rich alpha-2-glycoprotein repeat homology  
 C/Keywords: CAMP biosynthesis; duplication; phosphorus-oxygen lyase  
 F/1006-1029/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR>  
 F/1479-1564/Domain: yeast adenyLate cyclase catalytic domain homology <YACC>

Query Match 3.4%; Score 106.5; DB 1; Length 1839;  
 Best Local Similarity 21.3%; Pred. No. 22;  
 Matches 132; Conservative 75; Mismatches 220; Indels 193; Gaps 33;

QY 56 KEKTLILLDVSTK---NPATVNE--NPLSLQDPSITIHQWDFUSKRLVTLARGASP 110  
 DB 724 KKLQQLQDLSISKRFVYPEVINSCTLLQIDISYNTKHS--LP-VSINDVTLAK---- 776  
 QY 111 APLRFGGKRT---DPLQPNIR-----NPASRGPGP-----DYLYKRYEDDIVR 153  
 DB 777 -MNFNRRNLTSVGLSQMKRLRLNCRVTSIECHAPNOMFLTDNRISTFDDTLR 835  
 QY 154 SDVALDKQCKIAQHPDVMVLQREKAAQMHVLYLKEQPSNTYSNLLITARSIDKLVN 213  
 DB 836 ---LRTLEQNPITTSNVCQGNVWANTSLINKAKLSFS---AELSKLPR- 882  
 QY 214 ADCSGMLIFALNALRNPNN-----SNSSALSILKYSASKKYNISWELGNE----- 261  
 DB 883 ---LEKLEENNNLTQLPEINLTLIYLSVARNKLRSIPPEISDLRLSKS 931  
 QY 262 ---EPNNTRYMGRVAVNSQKDIYQLKSL-----QPIRYSRASLIGPNIGRRKN 312  
 DB 932 LDLSNNLRMLMN---NLEDLELTSLVNSSLTLGFHSGPAKFPASPS---PKLAK--- 981  
 QY 313 VTALIDGFMAKYSTVDVATWQHOCYIDGRVVKVMDFKITRLLD-----TLSDQIRKIQV 367  
 DB 982 ---SLV--FLSVADNNLTDSIF-----LVNTQNLKTLNLSINNVEISD--LKIQNL 1028  
 QY 368 VNTYTPG-----KTIWLEGVYTTSA-----GNNNIS-- 394  
 DB 1029 TELYISGNNFTSLPGEAVOHLRSILKVLMLNKNKLSLPAELSQLSRSLVDVGSNQLKYN 1088  
 QY 395 -DSYAGGLVNT---LGMLANQIDIVIRHSFDDHGVNHLVDQNPPLPDVWLSILY 448  
 DB 1089 ISNHYVDNMNRNNDKLTNPSGNRRF--IKSLADPEBKNDLSD-----LGLL- 1135  
 QY 449 KRLIPKYLAVHAGLQRRPRGVIYRDKLRIYAH-----CTNHN 489  
 DB 1136 KQL---RYLGMDVTLKSKYPPDESVSIRLTTASMINGRVGVADLTGSDVCSRDVT 1192  
 QY 490 HNYVSGSTLPIFINHSRKKIKLAGITRDKLVQY-----LLQPIGE-----GLK 536  
 DB 1193 FERFGRDEECILCYDKENENASGHKISKIRIDYDKILIRLEKYGESDGIKRALR 1252  
 QY 537 SKSVOLNGO---PLVWVDG 553  
 DB 1253 YSFLOLNKEINGMLVSVEDG 1272

## RESULT 10

C88264  
 protein kin-15 [imported] - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 17-May-2002  
 C/Accession: C88264  
 R/Anonymous: The C. elegans Sequencing Consortium.  
 A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
 A/Reference number: A75000; MUID:99069613; PMID:9851916  
 A/Note: see websites genome.wustl.edu/sec/C.elegans/ and www.sanger.ac.uk/Projects/C.eleg  
 A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A/Accession: C88264  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-484 <STO>  
 A/Cross-references: GB:chr\_11; PIDN:CAB01648.1; PID:G3878745; GSPDB:GN00020  
 C/Genetics:  
 A/Gene: kin-15



```

Db      364 FKEKTLNLTEDPEVYLVVIDONTDSWEKSLDGNNOY-----ELIKYISYANKSA 417
Qy      164 CTIAOHPDVMVLQREKA-----ACQELVYLKEQFSNTYSNLTARSLDK 209
      418 CR--KH-----LIILNKALDISMDKRLMIOSELNV-LKKNFQMT-----DA 459
Db      210 LYNPADCSGLHLIIPALNLRPNNSWSSSALSILKISA--SKKNI SWELG-----260
Qy      460 EFQFIPCSGL-----LGSNLTENITTKS--KYKSEFDSINYVPEWEGEPFSQ 507
Db      261 -----NEENNRRTMGRAVNSQGLKQYIOLKSLQPIRISASLYPNIGRPRKNI 314
Qy      508 LILVLEHNNKKTETLEEPVGT-----ILQSVIOPFAELNYSV-----KV 550
Db      315 ALLDGFMYAGSTVDAVWQ--HCY-IDGRVVKVMDPIKRLDLTIS-----DQIRKIQ 365
Qy      551 LINSGYIQ--SGQTEIHTQDEPHYGIYSRMKNSKQILETWTKNISVGLNPDILEVLV 609
Db      366 KVVNTYTPGKIM--LEGVTTISAGCTNNLSDSYAGFLMTLTIGMLANQIDVIRHGF 423
Qy      610 KIHNTEDFTKQPHIRKGDIIHSRKTNTLSPVLP-----NITKULALR-----LIKSI 659
Db      424 FDHGYHMLVDONENPLPDYMLSLLYKRLIGPKVLAHVAG--LQKPRPGVIRDKRI 480
Qy      660 QTHALSDPVLGSE-----LLVYHNLTHNAVYKVIIGTNDISINPQSLIVEVEIE 712
Db      481 YAHCTVHHNNVYRSGITFLIN 503
Qy      713 PDFALNVDSKTYNNIVLTSID 735

```

## RESULT 13

```

C97322
Probable alpha-arabinofuranosidase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: C97322
R:Rolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-835 <KIR>
A:Cross-references: GB:AE001437; PIDN:AAK81366.1; PID:G15026526; GSPDS:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3436
C:Superfamily: Streptomyces alpha-L-arabinofuranosidase SC135.05C

```

```

Query Match      3.4%; Score 105; DB 2; Length 835;
Best Local Similarity 20.5%; Pred. No. 8.5;
Matches 99; Conservative 61; Mismatches 172; Indels 148; Gaps 24;

Qy      23 APGALYLALILHH-----SLSSQAGDRPLPVDRAAGLKEKTLIILNDISTKNPVTVE 76
Db      66 ADGGLYSELLONOSFEFKDLSMTVDK-----TGSTST--AEVTSKPLNSNT 114
Qy      77 NPLSLQDPSIINDGMLDPLSKRLVTLA--RGLSPAVLRFGKRTDPLQONLRNPAKR 135
Db      115 HYELNCP-----DNNSSLKLVNSGYKGITV--NNAKIDFYFR--ARVNGK--157
Qy      136 GGGPPDYLLKNYEDDIYRSDVALDKQK--CTIAQH-----PDVYLQREKAAQMH 186
Db      158 -GNKVTITQLEDENGNALISEDKTIGKINGQWKTEGHLRATKSTSNALAVSITGEAKMN 216
Qy      187 VL-----LKEQFSNTYSNLTARSLDKLYNFADCSGLH 220
Db      217 DMSVLPQDTWKRRKYLKRDVLKRLKPLRFPFGGCIIVGNSKEELIYNKADITIG--274
Qy      221 LIFALNLRNPN--NSW--NSSSALSLKYSASKKNI SWELGNEP-----NNYRTWEGRA 273

```

```

Db      275 -----NVEERKENTNMGYNQSYGLGYEY-----FQLEDIATVPVLNCGMTQAG 324
Qy      274 VNSQGLKDYIOLKSLQPIRISASLYGNIGRPRKNVIALLDGFMKVASGTVAVT- 332
Db      325 VNG-----VPNVAVP-----GPDIDPIQNAVLD-----VEANGDASTY 360
Qy      333 WQCTYIGRVYKVDPIKTLTLDT-----LSQIKIKQVNVNTYTPGKIMLEGVTTIS 386
Db      361 WGRKRISGHKKPPN-LKYVALIGNEQMPYHHRFEAFQVNLQCKP-----GITLIS 412
Qy      387 AGTNNLSDSYAGFLMTLTIGMLANQIDVIRHSEFDHGYHMLVDONENPLPDYML 445
Db      413 NAGTSPSGSTFDDMMNMIK-----EKAPTVVDEHYMSPDWFLS 452

```

## RESULT 14

```

S73548
MG422 homolog C12.orf839 - Mycoplasma pneumoniae (strain ATCC 29342)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73548
R:Himmelreich, R.; Halbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73548
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-839 <HIM>
A:Cross-references: EMBL:AE000022; GB:U00089; NID:G1673882; PIDN:AB95870.1; PID:G1673882
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
C:Superfamily: hypothetical protein MG422

```

```

Query Match      3.4%; Score 104.5; DB 2; Length 839;
Best Local Similarity 20.2%; Pred. No. 9.3;
Matches 109; Conservative 85; Mismatches 181; Indels 165; Gaps 27;

Qy      70 PVRATNENF--LSIQD-----PSII-HDGMULDPLSKRLVTL--ARGSPAFRFGGR 119
Db      374 PQOSLNMKTIQALISELEKTLTPYIVLTHDAVIGFIKRLAFLQRYAGLARC--EQNKL 431
Qy      120 TDFLOFQNLNRPASRGPGPDYILKN---YEDDIYRSDVALDKQKCKIAQHPDVMV 175
Db      432 VNELKQNVLTNQNGYKGEVSQSYVNVQKAVFGFIQTVKALNVTYK-LKHTLDPEFLMN 490
Qy      176 LQREKAAQMHVYLKEQFSNTYSNLTARSLDKLYNF--ADCSGLH-----220
Db      491 IVQERCF-----EQLITTEKLDWKYKELNQLYVCRALNMLKQTOHFTTKYQFI 542
Qy      221 --LIFALN-----ALRRPNNSWSSSALSILKYSASKKNI SWELGNEPNRY 267
Db      543 THGVVDFFVNGRQNOTQASLKAHLNDPNSPKKLVNKNVKNYKFEANL--HHPAYL 599
Qy      268 TWEGRAVNSQGLK-----DYQLKS-----LQPIIYRASLYGNIGRPR 310
Db      600 LIPNR--NATLTENTMTTITQNLKTLQKQRAHYHLLQDIDILIQ--WYKKEIKQKQ 655
Qy      311 KVVIALDGPMMKVASGTVAVWQHCYIDGIVKMDPELKTRLIDTLSDQIRKIQKYNT 370
Db      656 QQIKRLKNY-----GTIANL-----LNTQISKYNNV--683
Qy      371 YTPGKIMLEGVTTISAGCTNNLSDSYAGFLMTNT-----LGMLAN 412
Db      684 ----RKTFF--VSEECILNRLOASNDLHFNLLNANVNVISFLCKCRQNPCKLNTNAN 736
Qy      413 QGIDVIRHSEFDHGYHMLVDONENPL--PDVYLKELIGPKVLAHVAGLQKPR 469
Db      737 --LKWLDNTRFGKIPSMWIFSDLNKINTKQRFKLYLFLKLLHPQVLV-----784

```

QY 470 PGRVIRDKLRIYAHCTNNHNNVYRGSTITLFINLRSRKIK-LAGTLRDKLVHQYLLQ 528  
DB 785 -----DSFVNPF-----NKHITVNTFRG-----LLIAHQNOGIAVLFNPNHNLVQKDFFTQ 829

## RESULT 15

B34087  
hypothetical protein (L1H 3' region) - human  
C/Species: Homo sapiens (man)  
C/Date: 30-Mar-1990 #sequence\_revision 30-Mar-1990 #text\_change 30-Sep-1993  
C/Accession: B34087  
R/Scott, A.F.; Schmeckpeper, B.J.; Abdelrazik, M.; Comey, C.T.; O'Hara, B.; Rossiter, J.  
Genomics 1, 113-125, 1987  
A/Title: Origin of the human L1 elements: proposed progenitor genes deduced from a conserved  
A/Reference number: A34087; MUID:8085185; PMID:3692483  
A/Accession: B34087  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 1-1280 <SC0>  
C/Superfamily: pol polyprotein

Query Match 3.4%; Score 104.5; DB 2; Length 1280;

Best Local Similarity 19.4%; Pred. No. 18;  
Matches 127; Conservative 94; Mismatches 216; Indels 217; Gaps 34;

QY 50 DRAAGLKEKTLILDVSTKNPVRT-----VNEPFLIQDPSIHDGWLDFLSKRL 101  
DB 136 DLQRDLDSHTIINGDFNT--PLSTLRSTRQKNKDIOELN--SALHQ-----DL 182  
QY 102 VTLARGSPAFRFGCRTPDLOFQNLNPNKSRGSGPPDYLNKVEDDIVRSDVALDKQ 161  
DB 183 IDIYRTIHP-----KSTETTF-----SAPHITYSKIDHILSKALLSKC 222  
QY 162 KGCKI-----AQHPVNLVLOREKAAQM-----LVY-----LKEQFSN----- 195  
DB 223 KRTETIITNCISDSHAIQLERIKLTQNHSTWLNLLNDYVHNEMKAEIKNFEETN 282  
QY 196 -----TYSNLLTARSJDKLYNPADCGSLHIFALNALRNPNNSMNSSALSLLKYSAS 250  
DB 283 ENKDTTYQNLMDTFKAV-----CRGKFT--ALNNHNR-----KQERS 317  
QY 251 KKYNI SWELNGNEPNRYRTMGRVNSQLGKDYIOLKSL-----LQPIRIYSRSLYGP- 304  
DB 318 KIDPLTSQI--KELEKQOTSKASRQETIKIRALKEIETQKLOKIN--ESRSWFEKI 375  
QY 305 -NIGRPKNVIA-----LDDGPKVAGS--TVDAVWQHCHYIDGRVKNWDFIKTRLDDI 356  
DB 376 NKIDRPLARLIKKGKKNQIDSIKNDKGDITDPTETIQ-----TTIREYKHLVYANK 427  
QY 357 LSPQIRKIQKVVNTYT-----PGKIMLEGVVTTSAGTNNLSDSVAAQ 400  
DB 428 L-ENLEMDKFLDVTYIPLRNQEEVESLNRPITGSEI--EATINSLEPTKSPGDPFTAE 484  
QY 401 F-----LMLNTLGMLANOGIDVYIRHSPFD-----HGYNHLVDQNFNPLPDY 442  
DB 485 FYGRYKELVPFLKLPQSIKEGI---LPSFEASIIILPKGRDYYTKENFRPIS-- 539  
QY 443 WLSLRYRLIGPKVIA-----VHVAGLQKRPGRV-----IRDKLRIYAHCTNNHNNH 490  
DB 540 -LNMIDAKILN-KILANQIQCHIKKLIHQVGRIPGMQGMFNIRKSNINVIQHINRTXK 597  
QY 491 NYVRGSTITFINLHRSRKIKIAGTLRDKLVHGYLLQPYGQSL-----K 536  
DB 598 NHM-----IISID-----AEKAFDKIQOPFMLKTKNGIDGTYLKIRAIYDKP 642  
QY 537 SKSYVNGQPLVWVDDGTLPELKRPLRAGTIVIPVTMGFYVKNVNALACR 590  
DB 643 TANIILNGQ-----KLEAFPLKGTGRGCPISPLIFNIVLEVLARATR 685

Search completed: May 6, 2004, 13:46:13  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:41:21 ; Search time 18 Seconds

(without alignments)  
1712.530 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088

Sequence: 1 MRYLCAPPEAMPSSNSRPPA.....PTMGFYVKNVNALACRRY 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	4.0	2026	1 CYAA_YEAST	P08678 saccharomyc
2	115	3.7	923	1 YEL4_YEAST	P40017 saccharomyc
3	114.5	3.7	837	1 XYNZ_CLODM	P10478 clostridium
4	106.5	3.4	1839	1 CYAA_SACKL	P23466 saccharomyc
5	106	3.4	488	1 K115_CAEEL	P34891 caenorhabd
6	104.5	3.4	839	1 Y422_MYCPN	P75135 mycoplasma
7	104	3.4	6359	1 BACC_BACLI	O68008 b bacitraci
8	101	3.3	535	1 COX1_HANWI	P48868 hansenula w
9	101	3.3	779	1 LON_BRECH	P67772 brevibacill
10	101	3.3	2971	1 VXC9_CHRE	O32065 chlamydomon
11	100.5	3.3	630	1 GIDA_TREPA	O83094 treponema p
12	100.5	3.3	870	1 SUL2_HUMAN	O81wus homo sapien
13	100.5	3.3	878	1 MSH4_YEAST	P40965 saccharomyc
14	100.5	3.3	1259	1 LINI_HUMAN	P08547 homo sapien
15	100.5	3.3	3343	1 YOG7_CAEEL	P34616 caenorhabd
16	99	3.2	546	1 SYR_TREAC	O9167 thermoplasma
17	99	3.2	871	1 SCL1_YEAST	O06245 saccharomyc
18	98.5	3.2	836	1 GK1_MOUSE	O60934 mus musculu
19	98.5	3.2	1356	1 ROM2_YEAST	P51862 saccharomyc
20	98	3.2	629	1 T3MH_HAELN	P71366 haemophilus
21	98	3.2	866	1 RECE_ECOLI	P15032 escherichia
22	97	3.1	543	1 YDE3_SCHPO	Q10437 schistosach
23	97	3.1	897	1 DPO1_HELPJ	O92169 helicobacte
24	96.5	3.1	544	1 RGSE_RAT	O08773 rattus norv
25	96	3.1	4036	1 RSLT_DUGBY	O66431 dugbe virus
26	95.5	3.1	709	1 STLI_CAEEL	Q13376 caenorhabd
27	95.5	3.1	2136	1 YCF2_MARPO	P09975 marichantia
28	95	3.1	2492	1 TAL4_DICDI	P44633 dictyostela
29	94.5	3.1	695	1 NIC3_DROME	Q9VC27 drosophila
30	94.5	3.1	832	1 ANR3_HUMAN	P57078 homo sapien
31	94.5	3.1	870	1 SUL1_MOUSE	O81009 mus musculu
32	94.5	3.1	1235	1 KPB2_HUMAN	P60019 homo sapien
33	94.5	3.1	1235	1 KPB2_MOUSE	Q6bwj3 mus musculu

34	94	3.0	625	1 GIDA_IACIA	O9cej4 lactococcus
35	94	3.0	641	1 LIP_STAYH	P04635 staphylococ
36	94	3.0	839	1 N196_YEAST	P34077 saccharomyc
37	94	3.0	1517	1 RPOC_CAME	O9p130 campylobact
38	93.5	3.0	870	1 SUL1_RAT	Q9v160 rattus norv
39	93.5	3.0	1106	1 C1C2_RABIT	P13806 oryctolagus
40	93	3.0	1075	1 PST2_SCHPO	O13919 schistosach
41	93	3.0	1300	1 POL2_MOUSE	P11369 mus musculu
42	92	3.0	430	1 SUPA_BUCAI	P57240 buchnera ap
43	92	3.0	461	1 GP13_YEAST	P23263 saccharomyc
44	92	3.0	534	1 COX1_KLITA	P20366 kluyveromyc
45	92	3.0	773	1 HEXB_ALITO	P49007 alteromonas

## ALIGNMENTS

RESULT 1  
CYAA\_YEAST STANDARD; PRT; 2026 AA.  
ID CYAA\_YEAST  
AC P08678;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (adenyl[yl cyclase]).  
GN CYR1 OR CDC35 OR HSR1 OR SRA4 OR YJL005W OR J1401.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_Taxid=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86079531; PubMed=2934138;  
RA Kataoka T., Broek D., Wiggler M.;  
RT "DNA sequence and characterization of the S. cerevisiae gene encoding adenylate cyclase."  
RL Cell 43:493-505(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA To Van D., Perea J., Jacq C.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=8286C / FY1679;  
RA de Haan W., Smits P.H.M., Griwell L.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 1042-2026 FROM N.A.  
RX MEDLINE=88165073; PubMed=3327602;  
RA Masson P., Lenzen G., Jacquemin J.M., Danchin A.;  
RL "Yeast adenylate cyclase catalytic domain is carboxy terminal."  
RN Curr. Genet. 10:343-352(1986).  
RN [5]  
RP MUTAGENESIS OF THR-1651.  
RX MEDLINE=9112042; PubMed=1991451;  
RA Feger G., de Vendittis E., Vitelli A., Masturo P., Zahn R., Verotti A.C., Kayounis C., Pal G.P., Fasano O.;  
RL "Identification of regulatory residues of the yeast adenyl[yl cyclase."  
RN EMBO J. 10:349-358(1991).  
RP FUNCTION: Plays essential roles in regulation of cellular metabolism by catalyzing the synthesis of a second messenger, cAMP.  
CC -| CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.  
CC -| COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
CC -| ENZYME REGULATION: The presence of GTP-bound Ras2 protein is required in order to elicit a magnesium-dependent adenyl[yl cyclase activity.  
CC -| SIMILARITY: Belongs to the adenyl[yl cyclase class-3 family.  
CC -| SIMILARITY: Contains 21 leucine-rich (LRR) repeats.  
CC -| SIMILARITY: Contains 1 PP2C-like domain.  
CC -| SIMILARITY: Contains 1 Ras-associating domain.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M2057; AAA34549.1; -;  
 DR EMBL; Z49280; CAA89295.1; -;  
 DR EMBL; X87611; CAA60917.1; -;  
 DR EMBL; X03448; CAA27175.1; -;  
 DR PIR; S56776; OBYB;  
 DR GenBank; 141621; -;  
 DR SGD; S0003542; CYR1;  
 DR GO; GO:0005886; C:Plasma membrane; IDA;  
 DR GO; GO:0007126; P:metolysis; IMP;  
 DR GO; GO:0000074; P:regulation of cell cycle; IMP;  
 DR InterPro; IPR001054; G\_cyclase;  
 DR InterPro; IPR001611; LRR;  
 DR InterPro; IPR003591; LRR\_cyp;  
 DR InterPro; IPR001932; PP2C-like;  
 DR InterPro; IPR000159; RA\_domain;  
 DR Pfam; PF00211; Guanylate\_cyc; 1;  
 DR Pfam; PF00560; LRR; 15;  
 DR Pfam; PF00481; PP2C; 1;  
 DR SMART; SM00044; C1CC; 1;  
 DR SMART; SM00332; PP2C; 1;  
 DR SMART; SM00314; RA; 1;  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1;  
 DR PROSITE; PS50200; RA; 1;  
 DR Lyase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding;  
 KW Magnesium.  
 FT DOMAIN 676 755 PAS-ASSOCIATING.  
 FT REPEAT 815 838 LRR 1.  
 FT REPEAT 842 862 LRR 2.  
 FT REPEAT 863 885 LRR 3.  
 FT REPEAT 886 908 LRR 4.  
 FT REPEAT 910 931 LRR 5.  
 FT REPEAT 932 955 LRR 6.  
 FT REPEAT 957 976 LRR 7.  
 FT REPEAT 977 999 LRR 8.  
 FT REPEAT 1001 1016 LRR 9.  
 FT REPEAT 1017 1040 LRR 10.  
 FT REPEAT 1042 1062 LRR 11.  
 FT REPEAT 1063 1086 LRR 12.  
 FT REPEAT 1088 1109 LRR 13.  
 FT REPEAT 1110 1132 LRR 14.  
 FT REPEAT 1134 1156 LRR 15.  
 FT REPEAT 1188 1209 LRR 16.  
 FT REPEAT 1210 1232 LRR 17.  
 FT REPEAT 1233 1256 LRR 18.  
 FT REPEAT 1258 1280 LRR 19.  
 FT REPEAT 1285 1308 LRR 20.  
 FT REPEAT 1319 1344 LRR 21.  
 FT DOMAIN 1369 1625 PP2C-LIKE.  
 FT DOMAIN 1626 2026 CATALYTIC.  
 FT METAL 1673 1673 MAGNESIUM (BY SIMILARITY).  
 FT METAL 1716 1716 MAGNESIUM (BY SIMILARITY).  
 FT MUTAGEN 1651 1651 T->I: WEAK PAS-INDEPENDENT ACTIVITY.  
 FT MUTAGEN 1651 1651 T->D: ATTENUATION OF THE RESPONSE TO  
 FT MUTAGEN 1651 1651 PAS PROTEINS.  
 FT CONFLICT 262 262 S->L (IN REF. 1).  
 FT CONFLICT 548 548 S->L (IN REF. 1).  
 FT CONFLICT 592 592 D->H (IN REF. 1).  
 FT CONFLICT 709 709 R->I (IN REF. 1).  
 FT CONFLICT 962 962 L->P (IN REF. 1).  
 FT CONFLICT 1388 1388 E->S (IN REF. 1).  
 FT CONFLICT 1427 1427 E->D (IN REF. 1).  
 FT CONFLICT 1461 1461 A->S (IN REF. 1).  
 FT CONFLICT 1566 1566 A->S (IN REF. 1).  
 FT CONFLICT 1735 1735 V->G (IN REF. 1).

FT CONFLICT 1956 1956 I -> V (IN REF. 3).  
 FT CONFLICT 1996 1996 C -> P (IN REF. 1).  
 FT CONFLICT 2009 2026 NVDELQYKAKKQDST -> MLTNFYKALRTQRIYOLE  
 FT FT FCS (IN REF. 3).  
 SQ SEQUENCE 2026 AA; 227832 MW; BEAB7419AB6989D0 CRC64;  
 Query Match 4.0%; Score 123; DB 1; Length 2026;  
 Best Local Similarity 22.6%; Pred. No. 1.1; Indels 162; Gaps 33;  
 Matches 141; Conservative 83; Mismatches 237;  
 QY 37 LSSQAGRRPLPVRAAGLKE-KTLLIDVSTKNV---RTVENELSLQDPSTIHGW 92  
 DB 891 LNLQNEHESLP---AGFVELKNLQLLDDSSNKFMYPEVINVCNLLQIDISYNTQIS 946  
 QY 93 LDFLSKRLVTLAR-GLSPAFLRGGRTPF--LQFQNLK-----NPAKSGGPGP 140  
 DB 947 LP-OSTKTLVLAQKNSHNKNTFGDLSERTDRTLNKYNRISLTKTASLQNLFLT 1005  
 QY 141 DYLLAKNYEDDIVRSADVADKQKCKIAQHPDVMVLQREKAQW-HLVLLKEQPSNTYSN 199  
 DB 1006 DNRISNFPD-----TLPKRLALEIQENPITSISFDQFPKNTSLTLNKAQLSSIPGE 1058  
 QY 200 LILTRASLDKYNPADSGHLITALLARPNKSNSSALSILKYSAKKYNISMEI 259  
 DB 1059 LITKLSFLEKLE-----LNQNNLTRLPQEI-SKTLKLVLSVANKKLEYLPEL 1106  
 QY 260 GNEPNRYT--MHGRAVNSQLGQDYIQLKSLQPIRIYRSASV-----YGPNIQR 308  
 DB 1107 -SOLKSLRTLDLHNHNRIDPVDGMENELITSNLSNAFNSLSNGFYHNSYGSYLSK 1165  
 QY 309 PRKNVIALDQPMKYAGSTDAVTWQ--HCYIDGRVYKVMDFLTRLLDLSQIRIQK 366  
 DB 1166 ----SLW--FFLAADQFDDAMPPLNCFVN---LKVNLSTYNSFDV---SHWKLSS 1211  
 QY 367 VVNTYTPGKXIMLGGVTVTSAGTNNLSDSYAAGFLW--INTLQMLANQGDV-----V 418  
 DB 1212 ITELTLSSNKL-----TTLSGDT-----VLKNSLSLKTMLNQNQLSPALSLNL 1256  
 QY 419 IRHSFDFHGYNLVDQNPFLPDY-WLS--LLYKRLGP----- 454  
 DB 1257 SOLSVFDVGANQLKYNISNHYDMNMRNKKELKYNISGNRFEIKSFISHDIDADISDL 1316  
 QY 455 -----KTLAVHVAAGLQKRPGRVIRDKLRIYACHTNHNH-----NYV----- 493  
 DB 1317 TVLPQKLVGLMDVTLNTKYPDENVFRKLTTSIINGKRYGVAADTLGGRDVYSRDVT 1376  
 QY 494 ---RGSITLFLINHSRKK-----IKLAGTLNDK-LVHOYLLQPYQF----- 533  
 DB 1377 FEFRRGNDECLCLHDSKQONADYGNHISRIVDIYDKILIRQ--LERYGDETNDNIKT 1434  
 QY 534 GLKSKSVOLNQ---ELVMVWDG 553  
 DB 1435 ALRFSFLQNKELINGMLNSVDNG 1457  
 RESULT 2  
 YEL4 YEAST STANDARD; PRT; 923 AA.  
 AC P40017;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypochemical 103.3 kDa protein in PRO3-GCD11 intergenic region.  
 GN YER024W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OC NCBI\_TaxId=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=6286C / AB972;  
 RX MEDLINE=97313264; PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,



RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Huntke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Nameth A., Norgren R., Oeffner P.,  
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.,  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.",  
 RL Nature 387:78-81(1997).  
 CC -1- SIMILARITY: Belongs to the carnitine/choleline acetyltransferase  
 CC family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U18778; AAB64557.1; -.  
 CC PIR: S50482; S50482.  
 CC GerMOnline: 139104; -.  
 CC SCD: S0000826; YER024W.  
 CC GO: GO:0004092; F:carnitine O-acetyltransferase activity; IMP.  
 CC GO: GO:0006066; P:alcohol metabolism; IGI.  
 CC GO: GO:0009437; P:carnitine metabolism; IMP.  
 CC InterPro: IPR000542; Carn\_acyl\_trans.  
 CC Pfam: PF00755; Carn\_acyltransfer\_1; PALISE\_NBS.  
 CC PROSITE: PS00439; ACYLTRANSF\_C\_2; 1.  
 CC PROSITE: PS00440; ACYLTRANSF\_C\_2; 1.  
 CC Hypothetical protein; Transferase; Acyltransferase.  
 KM SEQUENCE 923 AA: 103333 MW: B59AB881491D68A7 CRG64;  
 SQ  
 Query Match 3.7%; Score 115; DB 1; Length 923;  
 Best Local Similarity 18.7%; Pred. No. 1.4; Indels 206; Gaps 27;  
 Matches 115; Conservative 79; Mismatches 216;  
 33 LHSLSQAGDRRLPYDRAAGLEKKTLLIDVSTKNPVTNENFLS--LQDPS--II 88  
 98 LHDINNQSTRE-----IQDDVLRNPFVLADDLAPNTQDRSAVIV 142  
 89 HDGWLDSLSEKRLVTLARG--LSPAFIRFGKRTDFIQFQNL--RNPAKSGG----- 137  
 143 H-----SARFISALKQDILLPDIATNGKPLSMAFLNFGTTRSVPFGSEVENFD 195  
 138 -----PGPDY-----YKNEYEDIVRSDDVALDKQKGGKIAQHPDVMVLQ 177  
 196 LNKPYTASDLEDPEYSSDEDNDSEPTQKDPDRKRKHEEDIPGNGITIKRHPDSKILL 255  
 178 REKAQWELVLKQFSTNTYSLILIRASLDKLYN--PADCGHLIFALNAL----- 228  
 256 ISRGQVYTLLEVLT-----DSTWKIIYTAALTTIFNHIITKSSGIEKSTALGSLSHSFRN 310  
 229 -----RNPNNWNSSSALSILKYASAKKYNIMWEGNEPNNRYTHGGAUVNGSQL 279  
 311 WKYARKLQKQYPELHRIDBALFVLVDSQSEETT--DQDDYADISQMFNRITTRDK 368  
 280 GKDYIQLKSLLOPIRIVSRASLYGPNIGRPKNVIALIDGFMKYGASTVDA---VTWQHC 336  
 369 KCTSANCK-----RVFYGTSIINSKGQVSCVSRWYDKQLQVAV--TADAKATVWDSP 420  
 337 YIDGRVYVWMDFLKRLIDLTSDQIRKQKQVNYTTPGKKIM-----LEGV 383  
 421 TCDQSV--VLAFTS---EITYESVLARADVAGDQFSLMPVNTQMDPETHKLMATYI 474  
 384 TTSAGTNNISDSYAAGFLMINTLGMLANQGDIVVIRHSPFGHYNLVQDNENPLDPY 443  
 475 SADGGSPEDIPK-----LVNKK-----IMWSFS----- 498  
 444 LSLIKKLIGPKVLAVHVAAGLQKRPGRVRLDKLRIYAHCTNHNHNHYRGSTTLFIIN 503  
 499 -----NINLTHVHLSYTK-----LADLISKY-----DIYASISPL----- 528

QY 504 LHRSKKIKIAGTNRDLVHQVY-----LCPYQGEGLK-SKSVQINGQPLV 548  
 DB 529 ---GRSAQRIGVQPDMDVQVALQIAHYALXGRWVGLFHVSTRGFRGNSSPINIOSA 585  
 QY 549 MVD-----DGT 554  
 DB 586 LLELCQLEFISSIDGT 601  
 RESULT 3  
 XYNZ CLOTM STANDARD; PRT; 837 AA.  
 AC P10478;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Endo-1,4-beta-xylanase Z precursor (EC 3.2.1.8) (Xylanase Z)  
 GN (1,4-beta-D-xylan xylanohydrolase Z).  
 OS XYNZ.  
 OS Clostridium thermocellum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OC NCBI\_taxid=1515;  
 RN (1)  
 RP -SEQUENCE FROM N.A.  
 RC STRAIN=NCIB 10682;  
 RX MEDLINE=89008072; PubMed=3139632;  
 RA Grepinet O., Chebrou M.-C., Beguin P.;  
 RT "Nucleotide sequence and deletion analysis of the xylanase gene  
 RT (xynZ) of Clostridium thermocellum.",  
 RL J. Bacteriol. 170:4582-4586(1988).  
 RN (2)  
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS) OF 515-837.  
 RC STRAIN=NCIB 10682;  
 RX MEDLINE=95393242; PubMed=7664125;  
 RA Dominguez R., Souchon H., Spinelli S., Dauber Z., Wilson K.S.,  
 RA Chauvaud S., Beguin P., Alzari P.M.;  
 RT "A common protein fold and similar active site in two distinct  
 RT families of beta-glycanases.",  
 RL Nat. Struct. Biol. 2:569-576(1995).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylansidic  
 CC linkages in xylans.  
 CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS  
 CC WELL AS IN OTHER C.THERMOCCELLUM CELLULOSE ENZYMES. THIS DOMAIN  
 CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.  
 CC -1- SIMILARITY: Belongs to cellulase family F (family 10 of glycosyl  
 CC hydrolases).  
 CC -1- SIMILARITY: Contains 1 xynZ-type cellulose-binding (CBD) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M22624; AAA23286.1; -.  
 CC PIR: A31842; A31842.  
 CC PDB: 1XYZ; 29-JAN-96.  
 CC PDB: 1JUF; 31-OCT-01.  
 CC PDB: 1J72; 27-MAR-02.  
 CC InterPro: IPR006584; CBD IV.  
 CC InterPro: IPR005084; CBM 6.  
 CC InterPro: IPR002105; Dockerin\_1.  
 CC InterPro: IPR002048; EF-hand.  
 CC InterPro: IPR000801; Esterase\_put.  
 CC InterPro: IPR008979; Gal\_bind\_like.  
 CC InterPro: IPR001000; Glyco\_hydro\_10.  
 CC InterPro: IPR000379; Ser\_estrs.  
 CC Pfam: PF00442; CBM 6; 1.  
 CC Pfam: PF00404; Dockerin\_1; 2.  
 CC Pfam: PF00756; Esterase\_1.

DR Pfam; PF00331; Glyco\_hydro\_10; 1.  
 DR PRINTS; PR00134; GLYHYDRLASE10.  
 DR SMART; SM00606; CBD IV; 1.  
 DR SMART; SM00633; Glyco\_10; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN 2.  
 DR PROSITE; PS00448; CLOS\_CELLOSOME\_RPT; 2.  
 DR PROSITE; PS00591; GLYCOSYL\_HYDROL\_F10; 1.  
 DR Xylan degradation; Hydrolase; Glycosidase; Repeat; Signal;  
 3D-structure.  
 FT SIGNAL 1 28  
 FT CRAIN 29 837  
 FT ACT\_SITE 645 645  
 FT ACT\_SITE 754 754  
 FT DOMAIN 328 416  
 FT DOMAIN 430 487  
 FT REPEAT 430 453  
 FT REPEAT 464 487  
 FT DISULFID 783 789  
 FT HELIX 518 524  
 FT TURN 525 526  
 FT STRAND 528 533  
 FT TURN 535 535  
 FT HELIX 536 539  
 FT TURN 540 549  
 FT HELIX 543 552  
 FT STRAND 553 558  
 FT TURN 561 562  
 FT HELIX 564 567  
 FT STRAND 570 570  
 FT TURN 571 572  
 FT STRAND 573 573  
 FT HELIX 577 588  
 FT TURN 589 590  
 FT STRAND 592 599  
 FT HELIX 606 609  
 FT TURN 610 610  
 FT HELIX 615 632  
 FT TURN 633 636  
 FT STRAND 639 644  
 FT STRAND 647 647  
 FT TURN 649 650  
 FT STRAND 654 654  
 FT HELIX 658 663  
 FT TURN 665 666  
 FT HELIX 667 678  
 FT TURN 680 681  
 FT STRAND 683 688  
 FT HELIX 696 710  
 FT TURN 711 712  
 FT STRAND 717 720  
 FT STRAND 723 725  
 FT HELIX 730 745  
 FT TURN 746 747  
 FT STRAND 749 760  
 FT TURN 761 762  
 FT HELIX 765 785  
 FT TURN 787 788  
 FT STRAND 789 794  
 FT TURN 799 799  
 FT HELIX 800 807  
 FT TURN 809 810  
 FT STRAND 811 812  
 FT STRAND 817 817  
 FT TURN 819 820  
 FT STRAND 823 823  
 FT HELIX 825 834  
 FT SEQUENCE 837 AA; 92262 MW; DD4C29F04D12B6CD CRC64;

Query Match 3.7%; Score 114.5; DB 1; Length 837;  
 Best Local Similarity 21.0%; Pred. NO. 1.4; Indels 107; Gaps 20;  
 Matches 77; Conservative 54; Mismatches 128;

QY 128 LRNPARG---GPGDYIYKNTEDIVASDVALDKKCKINQHPDVLVLOREKAQM 184  
 DB 518 LRDAVEARGIKIGTCVNPYPYNSDPYNS-----IIQREFSMVVCENE----- 561  
 QY 185 HLVLTKQFSNTYVSNLLTLARSIDKLNFPADCGSL-----HLIFALNALRNP-----NSW 236  
 DB 562 ---MKDPALOPRQNVDFSKG--DQLAFERKGMQKRGITLWNH---QNPMLNNGW 613  
 QY 237 NSSSALSILKXYSAS-----KRYNISWELNE-----PNRYRMGRVAVNGSLGKDYLQ 285  
 DB 614 NRDSILAVMKKHITFTWTHYKGVEMDVANECMDSDGNGLRSSIMRWV---IGDYLD 669  
 QY 286 LKSLQPRIRYSAS-----LVG-----PNIGRPRXVIALLOGPMKVASSTVAVVWQH 335  
 DB 670 YAFR-----YAREADPDALLFTVDVIEDLG--PKSNAVNMKSKVRGVPIIDGVGFQC 722  
 QY 336 CYIDGRVYKVDPLKRLDPLTSPQIRKIQV-----VNTYTPGKKIMLGAVTTSA 387  
 DB 723 HFING-----NSPEVIASIDQNIKRYAIGIVAFTEIDIRIPQS---ENPATAFQ 770  
 QY 388 GGTNNLSSTVAGFLMNTLGLANQSIDVIRISFPDH-----GYNH--LVQNF 436  
 DB 771 VQANNYKELMKI-----CLANPNCNFTVMWGFTDKYTWLPGTPGNGNPLIDSNY 821  
 QY 437 NPLPDY 442  
 DB 822 NPKPAY 827

## RESULT 4

CYAA\_SACKL ID CYAA\_SACKL STANDARD; PRT; 1839 AA.

AC P23466;  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (adenylate cyclase).  
 GN CYR1.  
 OS Saccharomyces kluyveri (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxId=4934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91323718; PubMed=1864503;  
 RA Young D., O'Neill K., Broek D., Wiegler M.;  
 RL "The adenylate cyclase-encoding gene from Saccharomyces kluyveri.";  
 CC Gene 102:129-132(1991).  
 CC -!- FUNCTION: Plays essential roles in regulation of cellular  
 CC metabolism by catalyzing the synthesis of a second messenger,  
 CC cAMP.  
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.  
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (by similarity).  
 CC -!- SIMILARITY: Belongs to the adenylate cyclase class-3 family.  
 CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.  
 CC -!- SIMILARITY: Contains 1 P2C-like domain.  
 CC -!- SIMILARITY: Contains 1 Ras-associating domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC  
 CC EMBL; X56042; CAA39513.1; -.  
 DR PIR; J01145; OYBYK.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003591; LRR\_Typ.  
 DR InterPro; IPR001992; P2C-like.

```

DR InterPro: IPR00159; RA_domain.
DR Pfam: PF00211; guanylate_cyc; 1.
DR Pfam: PF00560; LRR; 15.
DR Pfam: PF00481; PP2C; 1.
DR SMART: SM00044; CYCC; 1.
DR SMART: SM00369; LRR_TPR; 2.
DR SMART: SM00332; PP2C; 1.
DR SMART: SM00314; RA; 1.
DR PROSITE: PS0125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE: PS50200; RA; 1.
DR Lysase; Repeat; Leucine-rich repeat; CAMP biosynthesis; Metal-binding;
Mgnesium.
KM DOMAIN 494 574 PAS-ASSOCIATING.
FT REPEAT 632 655 LRR 1.
FT REPEAT 659 679 LRR 2.
FT REPEAT 680 702 LRR 3.
FT REPEAT 703 724 LRR 4.
FT REPEAT 726 748 LRR 5.
FT REPEAT 749 771 LRR 6.
FT REPEAT 773 793 LRR 7.
FT REPEAT 794 818 LRR 8.
FT REPEAT 820 833 LRR 9.
FT REPEAT 834 857 LRR 10.
FT REPEAT 859 879 LRR 11.
FT REPEAT 880 903 LRR 12.
FT REPEAT 905 926 LRR 13.
FT REPEAT 927 949 LRR 14.
FT REPEAT 951 974 LRR 15.
FT REPEAT 1004 1025 LRR 16.
FT REPEAT 1026 1048 LRR 17.
FT REPEAT 1050 1072 LRR 18.
FT REPEAT 1074 1096 LRR 19.
FT REPEAT 1101 1124 LRR 20.
FT REPEAT 1135 1160 LRR 21.
FT REPEAT 1185 1440 PP2C-LIKE.
FT DOMAIN 1441 1839 CATALYTIC.
FT METAL 1488 1488 MAGNESIUM (BY SIMILARITY).
FT METAL 1531 1531 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCB1F2733CB C64;

Query March 3.4%; Score 106.5; DB 1; Length 1839;
Best Local Similarity 21.3%; Pred. No. 16;
Matches 132; Conservative 75; Mismatches 220; Indels 193; Gaps 33;

QY 56 KEKTLILLDVSTK---NPRTVNE--NFLSLQDPSIIHDGMDLFLSKKLVTLARGLSP 110
DB KTKKVLQLLDISSNKFNVPYEVINACTMLQIDLSYKTKHS--LP-VSINQLVKLAK---- 776
QY 111 APLRFGGKRT---DFLQFQNR-----NPAKSRGGGP-----DYLNKVEDIVR 153
DB 777 -KNLFNRLTSVGLSQPKRLRLTLNRCNRVTSIECHAPMLQNLFLTDNRISTDDDLTR 835
QY 154 SDVALDKQCKGKIAQHPDVMVLVLOREKRAQWHLVLEQPSNTYSNLILTPASIDKLYNF 213
DB 836 -----LRTLELQNPITSMVCGANWAMNTSLNKKAKLSRPS-----AEILSKLFR- 882
QY 214 ADGCGHLIPLNLRKRNPN-----SNASSALSLTKYSASKKNISWELG----- 261
DB 883 -----LEKLELNENNLTLQPEINLTLIYLVARNLDSIPDISLRSLSKS 931
QY 262 -----EPNNYRTWGRAVNGSQKDYIOLKSL-----QPIRTYSRASLYGPNIGRPRKN 312
DB 932 LDHSSNNLRLMNM--NLEDLELSLNVSSNLLGFHSPAKKFPASP-----PKLAK---- 981
QY 313 VIALIDGFMKYGSTVDAVTVQHCYIDGRVYKWMDFLKTLLD-----TISDQIRKIQV 367
DB 982 --SLL--FLSYADNNLTDSIWP-----LVNTEQNKTNLNNYNNEVEISD--LKLQNL 1028
QY 368 VNTYTPG-----KTIWLEGVVTTSAG-----GTNNLS-- 394
DB 1029 TELYLSGNFTSLGSEAVQHRLSKVLTMLNKNKLSLSPALISQSLRSLVDVGSNQKYN 1088
QY 395 -DSYAAGLFWLNT-----LGMLANOGIDIVIRHSFPDHGNHLYDONFNPIDPYWLSLLY 448

```

```

DB 1089 ISNHYVDMNRKNNKDKLTFNSGNKRF--IKSLDDEGKNLSD-----LGLL- 1135
QY 449 KLLIGPKYLAHVAVAGLQKRPGRGVIRDKLRIVAH-----CTNHNH 489
DB 1136 KQL---RVLGMDVTLKTSKVPDESVSIRLRTTASMINGMKRYGAVDTLGGSDSVCSRDVT 1192
QY 490 HNYVAGSTLTPLINLRKRRKIKLAGIRDLVQY-----LLQPYGGE-----GLK 536
DB 1193 FERFRGDEDCICLYDCKNNASSGKHSKIRIDYDKIIRLEKYGESDGIKRALR 1252
QY 537 SKSVQANGC--"PLVMVDDG 553
DB 1253 YEFLOANENMGMLVSVSDG 1272

RESULT 5
ID K115 CAEEL STANDARD; PRT; 488 AA.
AC P34691;
DT 01-FEB-1994 (Rel. 28, Created)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
GN KIN-15 OR M176.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristcol N2;
RX MEDLINE=94019384; PubMed=84113302;
RA Morgan W.R., Greenwald I.;
RT "Two novel transmembrane protein tyrosine kinases expressed during
RT Caenorhabditis elegans hypodermal development.";
RL Mol. Cell. Biol. 13:7133-7143(1993).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristcol N2.
RA Wilkerson J.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: May be specifically involved in cell-cell interactions
CC - regulating cell fusions that generate the hypodermis during
CC - postembryonic development. It has a role in the development of the
CC - HYP7 hypodermal synyutium.
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC - tyrosine phosphate.
CC - TISSUE SPECIFICITY: Hypodermal cells.
CC - DEVELOPMENTAL STAGE: Expressed during hypodermal development.
CC - SIMILARITY: Belongs to the Tyr family of protein kinases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collection -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L03524; AAA28151.1; -
DR EMBL: 278412; CAB01648.1; ALT_INIT.
DR PIR: I44330; I44330.
DR HSSP: P11362; 1FGK.
DR WormRep: M176.6; CE12470.
DR InterPro: IPR000719; ProC_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; ProC_kinase; 2.
DR SMART: SM00219; TyrcK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

```

DR PROSITE, PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE, PS50011; PROTEIN KINASE DOM; 1.  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;  
 KW Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 488 RECEPTOR-LIKE TYROSINE-PROTEIN KINASE  
 FT DOMAIN 27 488  
 FT TRANSMEM 51 70 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 71 488 POTENTIAL.  
 FT INP BIND 144 458 CYTOPLASMIC (POTENTIAL).  
 FT BINDING 150 158 PROTEIN KINASE.  
 FT ACT SITE 183 183 ATP (BY SIMILARITY).  
 FT ACT SITE 319 319 ATP (BY SIMILARITY).  
 FT CAREGARD 25 25 BY SIMILARITY.  
 SQ SEQUENCE 488 AA; 56544 MW; F891B9BEA128977 CRC64; (POTENTIAL).

Query Match 3.4%; Score 106; DB 1; Length 488;  
 Best Local Similarity 20.7%; Pred. No. 2.7;  
 Matches 89; Conservative 47; Mismatches 150; Indels 144; Gaps 21;

QY 62 LLDVSTN--PV--RTVNEFLSLQDPSI-IHDGML---DFLSSKRLVTLARGLSPA 111  
 DB 106 LLDVSTNNEETPTVPKPPINERISNLEFPDFEIDQAKLEISBDKLGSGFGEVCGYGL--L 163  
 QY 112 FLRFGGKRTDPLQFQNLNPAKSGRGPDPYLYKNEYDDIVRSVALDQKQCKIAQHPD 171  
 DB 164 SMETSTNETDTL--QLTSVAVKQSNPTQENCKMIED-----ETKMCALGRNEN 212  
 QY 172 VMLVLRKAQOMHLVLKEQPSNTYSNLTILARSJDKLYNPADCGLHLIFALNALRN 231  
 DB 213 ILAIIGAVTN-----SGSARNLLIV-----EFVECG----- 239  
 QY 232 PNNSSNSALSLIKTSASKKYNISWELGNEPNRYTMRGPAVNSQLGKDYIQLKSLQ 291  
 DB 240 -----DLKFLKEKKSIFKDELYENKGY-----LL 265  
 QY 292 PIPIYRSLSYGNIGRPRKNTI-ALIDGPMKAGSTVDAVTWQHCYIDG----- 340  
 DB 266 PKSIIRKTYMFEN-----EDDVIEESIDSL-----CTSDLSFSYQIAEGMEYLAIPCV 316  
 QY 341 -----RYKQVMDPLKTRLDLTSLDQIRKIOKVNTYPRGKIMLEGVVTS 386  
 DB 317 HRDLARNVTLNOKTRIRIDFGLARKYQ--DGYRITKVGVTMPPAR--MAAPRYAR 372  
 QY 387 AGGTNNLSD--SYAAGFLWNTLQMLANQI---DVVIRHSFPGHYNHLDQNFNPLPD 441  
 DB 373 -GKTEKSDVMSYGVSLYEMFSLGELPYSNVNSDV-----FEHVQGNQLPMPQ 421  
 QY 442 YMLSLYKRL 451  
 DB 422 YCHPKMIDRM 431

RESULT 6  
 Y422 MYCPN STANDARD; PRT; 839 AA.

ID AC P51575;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypoethetical protein MG422 homolog (Cl2\_crf839).  
 GN MN620 OR MP222.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelfreuch R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
 RA Hermann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma

RT pneumoniae";  
 RL Nucleic Acids Res. 24:4420-4449 (1996).  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000022; AAB95870.1; -  
 DR PIR; S73548; S73548.  
 KW Hypoethetical protein; Complete proteome.  
 SQ SEQUENCE 839 AA; 99912 MW; 3E3523E18BCDFECC CRC64;

Query Match 3.4%; Score 104.5; DB 1; Length 839;  
 Best Local Similarity 20.2%; Pred. No. 7.5;  
 Matches 109; Conservative 85; Mismatches 181; Indels 165; Gaps 27;

QY 70 PRTVNEF--LSQD-----PSTI-HDGMFLDSSRLVTL--ARGSPALFRGGKR 119  
 DB 374 PGLSLNMYQATISFELKTRPVLTHDAVIGFLISRLFLQRYAKLNRK--EQHNKL 431  
 QY 120 TDFLOPQNLNPAKSGRGPDPYLYKNEYDDIVRSVALDQKQCKIAQHPDMLV 175  
 DB 432 VNELKQVNLNQNYKGEVSQSVVQKAVFGFIQTVKALNLYTK-LKHTLDPEFLMN 490  
 QY 176 LQREKAQOMHLVLKEQPSNTYSNLTILARSJDKLYNF--ADCSGLH----- 220  
 DB 491 IVQRCF-----EQLTYEKLDTKYNQLVFCRALWNTLHKQTQHFYTKQFI 542  
 QY 221 --LIFALN-----ALRRPNNSNSASSALSLIKTSASKKYNISWELGNEPNRY 267  
 DB 543 THGVVDVFQNGRNQTPASLKNLNRDNMSPPKLLVNTKVFYFANL--HHQAYL 599  
 QY 268 TMAGRAVNSQLK-----DYQLKS-----LLQPIRYSRSLYGNIGRPR 310  
 DB 600 LLEPR--NATLTLEANTTTINQLKTKQKMPRAHYLLQDRILO--WLYKEIKQKQ 655  
 QY 311 KNVIALDGEFMKAGSTVDAVTWQHCYIDGRVYKMDPLKTRLDLTSLDQIRKIOKVNT 370  
 DB 656 QQIKALKNY-----GTLNKL-----LNTQISKVNNVV-- 683  
 QY 371 YTPGKRTWLBGVVTSAGGTNNLSYAGFLWNT-----LQMLAN 412  
 DB 684 ---RKTFF--VDSECDNRLQASNKLFHNLNMAVNIISCLKKCRQNPCKLRTAN 736  
 QY 413 QGIDVIRHSFPGHYNHLDQNFNPL--PDYWLSTLYRLLGPKYLAHVAVAGLQKPR 469  
 DB 737 --LQMLDNTKQNIPEWMLFSDLNKNTQKQRFYLLFGLLHPQVLV----- 784  
 QY 470 PGRVIRDKLRIYAHCTNHNHNHYVRSITLFIINLHRSRKIK-LAGTLRDKLVHQLQ 528  
 DB 785 -----DSFVNF-----NKHTYNFTRG-----LLIHAQNOGIAVLENDPNNHNVKQFFTQ 829

## RESULT 7

BACC\_BACLI STANDARD; PRT; 6359 AA.

ID AC O68008;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Bacitracin synthetase 3 (Bac3) [includes: ATP-dependent isoleucine  
 DE adenylation (IleA) (Isoleucine activase); ATP-dependent D-phenylalanine  
 DE adenylation (D-PheA) (D-phenylalanine activase); ATP-dependent histidine  
 DE adenylation (HisA) (Histidine activase); ATP-dependent D-aspartate  
 DE adenylation (D-AspA) (D-aspartate activase); ATP-dependent asparagine  
 DE adenylation (AsnA) (Asparagine activase); Aspartate racemase  
 DE (EC 5.1.1.13); Phenylalanine racemase [ATP hydrolyzing]  
 DE (EC 5.1.1.11)].  
 GN BACC.



```

CC      1-3form the functional core of the enzyme complex. CO 1 is the
CC      catalytic subunit of the enzyme. Electrons originating in
CC      cytochrome c are transferred via the copper A center of subunit 2
CC      and heme A of subunit 1 to the bimetallic center formed by heme A3
CC      and copper B.
CC      -1- CYTOCHROME C ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC      c + 2 H(2)O.
CC      -1- PATHWAY: Respiratory chain; terminal step.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC      inner membrane. Contains 12 potential transmembrane domains.
CC      -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; D31785; BAA0563.2; -.
CC      DR      PIR; S58740; S58740.
CC      DR      HSSP; P98002; 1AR1.
CC      DR      InterPro; IPR000883; COX1.
CC      DR      Pfam; PF00115; COX1; 1.
CC      DR      PRINTS; PR01165; CYCOXIDASE1.
CC      DR      PROSITE; PS00077; COX1; 1.
CC      KW      Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
CC      Respiratory chain; Inner membrane.
CC      FT      METAL 63
CC      FT      METAL 242
CC      FT      METAL 246
CC      FT      METAL 291
CC      FT      METAL 292
CC      FT      METAL 377
CC      FT      METAL 379
CC      FT      METAL 426
CC      FT      CROSSLINK 242
CC      FT      SEQUENCE 535 AA; 59025 MW; 9767C4EFAD1AD50A CRC64;
CC      (by similarity).
CC      Query Match 3.3%; Score 101; DB 1; Length 535;
CC      Best Local Similarity 22.2%; Pred. No. 7.3;
CC      Matches 68; Conservative 46; Mismatches 102; Indels 90; Gaps 19
CC      -----
CC      293 IRIYRSLIYGNIGRPRKQNVLTALIDGFMKAVGASTVDATVW-QRCYIDGRVYKVMDFLKT 351
CC      260 VSTYSKKKEVFGF-----ISMYAASIA-FLGFLVSHHWIYV-----DD-ADT 302
CC      352 RLIDTLSDQIRIKQVNTYTPGKKIWEAGVTTAGAGTNNSDS--YVAGELMTNTLG- 408
CC      303 RAYFSSNTMVAIVPGIKIFS-----NL-----ATLYGGSIRLAVMLYALALFLFTICG 353
CC      409 ---MLANOGIDVIRHSFFDHQYNLV--DQNFNPRLDYV-----LSLIYKR----- 450
CC      354 LTGVALLANASLDVAHFDTYYVGHFFYVLSMGAFSLFAGYYVWSPQILGLFENRILAQI 413
CC      451 -----LIGKY--LAVHVAGLGRKPRPGVINDKRLRYAHGCHNNHHNVVVG----- 495
CC      414 QEWLITFGVNAVTFMVFHFLGLQMPR-----RIPIYPAIYAGMNVVSSIGSVIAT 464
CC      496 SITFLINLHRSRKIKIAGTLRDKLVHQ--YLAOP-----YGOGLSKSVQ--L 542
CC      465 SLALFYIYID-----QLINGLTNKKIDKNSVYSKAPDEVESNTIIPANNSIKSASIEFL 519
CC      543 NGQPLV 548
CC      520 NSPPAI 525
CC      Db
CC      RESULT 9
CC      ID _ION_BRECH STANDARD; PRT; 779 AA.
CC      AC P36772;

```

[illegible]

```

Db      425 DKLASFRGDPASALTEVLDPNQNDKFSDH-----XIEETYLITNM 466
Qy      200 LITLAPSLKLVNFAOCGLHILFALNALRRNPNNSWSSALSLKYSKKYNI--SW 257
Db      467 FITTANSLDTI-----PRLLDMEVVISIGYTELEKNTILRGY 505
Qy      258 ELGNEPNRYTRMGRAVNSQLGKDYIOLK--SLLOPIRIYRASLYG--PNIGRPKNYI 314
Db      506 LEPKQMED--HG-----LQKDKLQNMEDAMKLVLTYREA--GVANINREAAVNC 552
Qy      315 -----ALLDQFMK--VAGSTVNAVWQHCYIDGRVYVMDFLKRLDLDQIRKIOK 366
Db      553 RKAATIVGGEKKRVVVTATLLEALGKPRYGLAEK-----KQVGSV-- 597
Qy      367 VVNTYTPGKKIWLEGVTTVSAGG--TNNLSDSYAAGFLMNTLGLMLAN-----QGIDVVI 419
Db      598 -----TGLAWTQAGDGLNVEVSLAGKGLTLTGQDGVKESQAALFSYI 644
Qy      420 RHSEFDFHGNHLYDQNPENLPDYWLSLYKRLIGPKYLAHV--AGLQKRPGRVIRDKL 478
Db      645 RSPASEWG--IDPEFEKND-----IHIVEGAIPIKDGSPAGITMAT 685
Qy      479 RIRAHCTNHNHNYVRSGITFLIINLRSRKIKLAG--TLRDKLVHGYLLQPYGQGLK 536
Db      686 ALVSALT-----GIPV-----KKEVGMGTGELTRGRVL-----PIG--GLK 719
Qy      537 SKSVQLNGQPLVMVDGTLPELKPRLPAGRTLVIRP 573
Db      720 EKCMASAH-----RAGLITITILP 736

```

## RESULT 10

```

ICX3_CHLRE STANDARD; PRT; 2971 AA.
ID YCX9_CHLRE
AC Q32065; Q95635;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 341.7 kDa protein in pebd-pbdc intergenic region
DE (ORF2971) (ORF8).
OS Chlamydomonas reinhardtii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137c / CC-125;
RA Watson A.T., Puton S.;
RT "Unidentified open reading frame ORF2971 (ORF8) from the chloroplast
RL genome of Chlamydomonas reinhardtii."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 2635-2971 FROM N.A.
RC STRAIN=137c / CC-125;
RX MEDLINE=93305500; PubMed=2663467;
RA Rochaix J.D., Kuchka M., Mayfield S., Schirmer Rahire M.,
RA Girard Bascou J., Benoun P.;
RT "Nuclear and chloroplast mutations affect the synthesis or stability
RL of the chloroplast psbc gene product in chlamydomonas reinhardtii."
RL EMBO J. 8:1013-1021(1989).
RN [3]
RP COMPLETE PLASTID GENOME.
RX MEDLINE=2305394; PubMed=12417694;
RA Maul J.E., Lilly J.W., Cul L., dePamphilis C.W., Miller W.,
RA Harris E.H., Stern D.B.;
RT "The Chlamydomonas reinhardtii plastid chromosome: islands of genes in
RL a sea of repeats."
RL Plant Cell 14:2659-2679 (2002).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U62843; AAB05800.1; -
DR EMBL; X13879; CA32083.1; -
DR EMBL; BK000554; DAA00965.1; -
DR InterPro; IPR003959; AAA_Atpase_centrl.
DR Pfam; PF00004; AAA; 1.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 2971 AA; 341663 MW; 3BB294AF2248348A CRC64;

```

Query Match 3.3%; Score 101; DB 1; Length 2971;  
 Best Local Similarity 20.7%; Pred. No. 81;  
 Matches 127; Conservative 87; Mismatches 207; Indels 192; Gaps 34;

```

Qy      62 LLDVSTKQPVRYNE-NPLSLQDLPSTIHQWLDPLSSKRLVT---LARGSPAFRFG 116
Db      306 LLD--SKAATQPSLSLVFTQNLQKTFSLWILPVAGLALITPTLLTLTGQSVQKFN 363
Qy      117 ---GKIDFLQFQNLNPAKSRG-----GPGDYLYKNY-----EDIVRSQVALDKQ 161
Db      364 SFINKKTDMVLSNTEPKSFGTPTLFGTSVEIYLPNSYWPKEGSGGINRVNNSINAV 423
Qy      162 KGCKIACHPDVWLVLOREK-----AAQMHVLYLKEQPSNTYSNLI--LTARSLDLQYFAD 215
Db      424 KKNATYAN--LVLDSESGEVATSPONDLSICVNNLNYNSKNTALSTKRLFLFSA 479
Qy      216 CSGL-----HLIFALNALRRN--PNNSWSSSALSGLK 246
Db      480 IKSNAIKHKTQSPFSVENTTLLGNNSPVYKGFKSSINAFSSYLPSTNVHSMPLTSLP 539
Qy      247 Y--SASKKYN-----ISWEIGNEP-----NRYRTMG--RAVNGS 277
Db      540 YLKAIISPLYSKFMIDHSIKFTPTTKLQHKLNKSPKQNTYKTQNFQGLDALNSF 599
Qy      278 QLGKDYIQ---LKSLLQPIRIYERASLYGNIGRPKNYIALDGEPMKV-----AGS 326
Db      600 SRGQVNFRTNHFILNSNRPLAHNYQA-----LKLINGEYQYQNNLQINCNK 645
Qy      327 TVDAVWQHCYIDGRVYVMDFLKRLDLDTSQIRKIQKRVNTYTGOKKIWLEGVTT 386
Db      646 TLD-----LNRKNKLVYGVHKSHEFNQKCSQIVYKQSLYN-----BDLCIR 687
Qy      387 AGGTNNLSDSYAAGFLMNTLGLMLANGIDVIVIRHSFFDGHYNLVD---QNFNPLPDY 442
Db      688 GNGT-KVVDYSHGDKLSNKGIVLDY---FVGNLLEFNKNTNTIINKDGQNTXK-- 741
Qy      443 WLSLI---YKRLIGPYLAHV--VAGLQKRPGRVIRDKLR1-YAHCYNH---HNH 490
Db      742 -LNLKTTVPFKTIL-KKQTSINSIVANEQ-----TRNNLWGIHFNGHLSVSNNA 791
Qy      491 NYVRGSITLFIINLRSRKIKLAGTLEDKLVHGYLLQPYGQGLKSKSVQALNQCPVMV 550
Db      792 NLTG-----RPVKFIYKRDKRLNSYLI--YNNQMLK-KFIDLNNN----- 830
Qy      551 DDGTLPELKPRL 563
Db      831 -----FLKPKPL 837

```

## RESULT 11

```

ID GIDA_TREP8 STANDARD; PRT; 630 AA.
AC G83084;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE Glucose inhibited division protein A.
GN GIDA OR TP0044.
OS Treponema pallidum.

```

CC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sedgwick E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Arriach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatcher B., Horst K., Roberts K., Sandusky M., Weisman J., Smith H.O.,  
 RA Venter J.C.,  
 RA "Complete genome sequence of Treponema pallidum, the syphilis  
 RT Spirochete."  
 RL Science 281:375-388(1998).  
 CC -1- FUNCTION: Not known.  
 CC -1- SIMILARITY: Belongs to the gida family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, AE001189; AAC65038.1; -  
 DR F1R; C71374; C71374.  
 DR TIGR; TP0044; -  
 DR HAMAP; MF 00129; -1.  
 DR InterPro; IPR001327; FAD pyr\_redox.  
 DR InterPro; IPR002218; GIDA.  
 DR Pfam; PF01134; GIDA\_sub.  
 DR PRINTS; PR00368; FADPVR.  
 DR ProDom; PD003738; GIDA\_1.  
 DR TIGRFAMs; TIGR00136; gida\_1.  
 DR PROSITE; PS01280; GIDA\_1; 1.  
 DR PROSITE; PS01281; GIDA\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 630 AA; 70175 MW; 1B526F02C1EC275 CRC64;  
 Query Match 3.3%; Score 100.5; DB 1; Length 630;  
 Best Local Similarity 21.3%; Pred. No. 9.9;  
 Matches 77; Conservative 53; Mismatches 128; Indels 103; Gaps 20;  
 QY 276 GSQUGK-DYIQKSLQPIRISRLVGNPGRKRVIALLDPMFMRAGSTYDAVW 333  
 DB 68 GSENGKPADCMQO---YRLNKR--GPAVQAPRIQDKFL--YAKVKYTLCTQH 118  
 QY 334 QHCYIDRVVAVKDFLRLDLTSDQIRIKQVNTYTPGKKIMLEGVTTTSGATNLT 393  
 DB 119 LHYQD---TVDVCSNTTDAGVAYGAHAHV--ARGRISARAVLVLTG----- 166  
 QY 394 SDVYAGFLM-----NTLGMLANQGDVVRHSFPDNG-----YHLYDQNF 436  
 DB 167 --FMEBRVITGTEYAEGRIGSHABGLAARKKGFQNGRIKGTPTARVLRKSYD-- 221  
 QY 437 NPLPDVWLSLYR---LIGPKVAVVAVAGLQRRKPRGVRIDKLRIYAHCTNNHNNY 492  
 DB 222 -----LSVMEKQEDAIRPSPFA-HV-----EINRPHADCTINTNETHQL 263  
 QY 493 VRGSITLFINHRS---KKIKLAGT-----LRDGLVQYIOLPFGDGSLKS 537  
 DB 264 IRE-----NFRSPFSGRIKAVGTRVCPSTEDKRVKRPDPDIRQLYIEP---EGDIT 313  
 QY 538 KSVQLANGQPLVWVD-----GLPEIK---RPLRAGTLYIVPTVAGFYV-VGVNVA 586  
 DB 314 EELYLINGLSCLPEDIQDEMIRITIPGMRVAVTRPAAVAVYALFPVQGIQLQTRVSG 373  
 QY 587 L 587

DB 374 L 374  
 RESULT 12  
 ID SUT2\_HUMAN STANDARD; PRT; 870 AA.  
 AC Q81WU5; Q96SG2; Q9H1H0; Q9URJ3; Q9ULH3;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Extracellular sulfatase Sulf-2 precursor (EC 3.1.6.-) (HSulf-2).  
 GN SUT2 OR KIAA1247.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF  
 RP 88-CYS-CVS-89.  
 RC TISSUE=Lung;  
 RX MEDLINE=22370956; PubMed=12368295;  
 RA Morimoto-Tomita M., Uchimura K., Ward Z., Hemmerich S., Rosen S.D.,  
 RT "Cloning and characterization of two extracellular heparin-degrading  
 RT endosulfatases in mice and humans."  
 RL J. Biol. Chem. 277:49175-49185(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,  
 RA Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XV.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 6:337-345(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida K.N., Babage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.K., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,  
 RA Clegg S., Coley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehaeaelaio M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Patrington S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.B., Sehra H.K., Showstee R., Sims S.,  
 RA Skuce C.D., Smith W.L., Sodeitund C., Stewart C.A., Sulston U.B.,  
 RA Swann R.M., Symmons N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20."  
 RL Nature 414:865-871(2001).  
 RN [4]  
 RP SEQUENCE OF 566-870 FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,



RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stalderon M., Soares M.B., Ronaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,  
RA Rana S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., Mowbray P.J., McKernan K.U., Malek U.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S.A., Gay L.J., Hultyk S.W.,  
RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy U., Helton E., Ketterman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Sheychenko Y., Bouffard G.G.,  
RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz U., Myers R.M.,  
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: Exhibits arylsulphatase activity and highly specific  
CC endoglucosaminase-6-sulphatase activity. It can remove sulfate from  
CC the C-6 position of glucosamine within specific subregions of  
CC intact heparin.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum and Golgi stack.  
CC -1- Also localized on the cell surface (By similarity).  
CC -1- TISSUE SPECIFICITY: Expressed at highest levels in the ovary,  
CC skeletal muscle, stomach, brain, uterus, heart, kidney and  
CC placenta.  
CC -1- MISCELLANEOUS: Shows maximal activity at pH 7.0 and 8.0.  
CC -1- SIMILARITY: Belongs to the sulfatase family.  
CC -1- CAUTION: Ref.3 (CAC17694) sequence differs from that shown due to  
CC erroneous gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AY101176; AAM76861.1; -; ALT INIT.  
CC EMBL; AB033073; BAA86561.2; -; ALT INIT.  
CC EMBL; AL354813; CAC39100.1; -;  
CC EMBL; AL133001; CAB61349.1; ALT INIT.  
CC EMBL; AL034418; CAC17694.1; ALT SEQ.  
CC EMBL; BC020962; AAM20962.1; ALT INIT.  
CC InterPro: IPR000917, Sulfatase.  
CC Pfam: PF00884, Sulfatase; 1.  
CC PROSITE: PS00523, SULFATASE\_1; 1.  
CC PROSITE: PS00149, SULFATASE\_2; FALSE NEG.  
CC HydroLase; Signal; Glycoprotein; Endoplasmic reticulum; Golgi stack.  
CC FT SIGNAL 1 24  
CC FT CHAIN 25 870  
CC FT MOD\_RES 88 88  
CC FT  
CC CARBOHYD 65 65  
CC CARBOHYD 112 112  
CC CARBOHYD 132 132  
CC CARBOHYD 149 149  
CC CARBOHYD 171 171  
CC CARBOHYD 198 198  
CC CARBOHYD 241 241  
CC CARBOHYD 561 561  
CC CARBOHYD 608 608  
CC CARBOHYD 717 717  
CC CARBOHYD 754 754  
CC CARBOHYD 764 764  
CC COMPLET 88 89  
CC CONFLICT 76 76  
CC SEQUENCE 870 AA; 100454 MW; 74B1069CE2774D73 CR664;  
CC  
CC Query Match 3.3%; Score 100.5; DB 1; Length 870;  
CC Best Local Similarity 19.3%; Pred. No. 16;  
CC Matches 65; Conservative 50; Mismatches 100; Indels 121; Gaps 18;

QY 86 SIHGWLDLSSKRLVTLARGLSPALRFGKRTDPLQFQNLNPAKSGGPGDYUK 145  
Db 151 SYVPGKKEV-----GLT-----KNSPFYVTLGRNGVKEHGSD---YAK 189  
QY 146 NYEDDIYSVALDKQCKCIAHPDVMYLQR-----EKAQMLVLLKPFSTNYEN 199  
Db 190 DYLLDLITNDVSVEFFRISKQVPHRPVLMWISHAHPGPDSDA-----PQYSRLEPN 241  
QY 200 LILFARSLDLYNPADSGHLIFALNALRPNNSWSSSALSLKYSASK-----252  
Db 242 ---ASGHITSYVAP-----NDKEM-----IKRYGPMKIHMEFT 276  
QY 253 -----NISWELGNPNVNTMGRANVNSQLGQDYIQLKSLLQ 291  
Db 277 NMLQRKRLQTLMSVDSMETIYVMTLVETGELDNTIYV--TADHGHIQO-FGLVYGRKSW 333  
QY 292 PIRIYSRASY--GPN--GRPRKNVALLD---GPMKVGSTVDVNTWCHCYIDRKYK 344  
Db 334 PYEDIVPVPYVROPVNAAGCLAPHIVNLIDLAPITIDIGLDIPA-----DMDGKST 386  
QY 345 VMDFLKRLDLYNPADSDQIRKIQKVVNTYTPGKK--IW 378  
Db 387 -----LKLDDT-----ERPYNREHLKKGRVW 408  
  
CC RESULT 13  
CC MSH4 YEAST STANDARD; PRT; 878 AA.  
CC AC P40965;  
CC DT 01-FEB-1995 (Rel. 31, Created)  
CC DT 01-FEB-1995 (Rel. 31, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE MSH4 OR YFL003C.  
CC GN MSH4 OR YFL003C.  
CC OS Saccharomyces cerevisiae (Baker's yeast).  
CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
CC NCBI\_TaxID=4932;  
CC  
CC [1]  
CC SEQUENCE FROM N.A.  
CC MEDLINE=95094270; PubMed=8001134;  
CC RA Rose-Macdonald P., Roeder G.S.;  
CC RT "Mutation of a meiosis-specific Muts homolog decreases crossing over  
CC but not mismatch correction.";  
CC RL Cell 79:1069-1080(1994).  
CC  
CC [2]  
CC SEQUENCE FROM N.A.  
CC STRAIN=S288C / AB972;  
CC RX MEDLINE=95400292; PubMed=7670463;  
CC RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
CC RA Sasamura S.-I., Sasamura M., Tsuchiya Y., Soeda E., Yokoyama K.,  
CC RA Yamazaki M., Tashiro H., Eki T.;  
CC RT "Analysis of the nucleotide sequence of chromosome VI from  
CC Saccharomyces cerevisiae.";  
CC RL Nat. Genet. 10:261-268(1995).  
CC  
CC [3]  
CC SEQUENCE FROM N.A.  
CC STRAIN=S288C / AB972;  
CC RX MEDLINE=96381249; PubMed=8789262;  
CC RA Naitou M., Ozawa M., Sasamura S.-I., Kobayashi M., Hagiwara H.,  
CC RA Shibata T., Hanaoka F., Watanabe K., Ono A., Yamazaki M., Tashiro H.,  
CC RA Eki T., Murakami Y.;  
CC RT "Sequencing of a 23 kb fragment from Saccharomyces cerevisiae  
CC chromosome VI.";  
CC RL Yeast 12:77-84(1996).  
CC  
CC -1- FUNCTION: Involved in meiotic recombination. Facilitate crossovers  
CC between homologs during meiosis.  
CC -1- SUBUNIT: Heterooligomer of MSH4 and MSH5.  
CC -1- SIMILARITY: Belongs to the DNA mismatch repair muts family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U13999; AAA6234.1; -  
 DR EMBL: D50617; BAA09235.1; -  
 DR PIR: A55201; A55201.  
 DR GerMOnline; A55201.  
 DR SGD: S0001891; MSH4.  
 DR GO: GO:0000228; C:nuclear chromosome; IDA.  
 DR GO: GO:0003677; F:DNA binding; IPI.  
 DR GO: GO:0007131; P:meiotic recombination; IMP.  
 DR InterPro: IPR000432; Muts\_C.  
 DR InterPro: IPR007860; Muts\_III.  
 DR InterPro: IPR007696; Muts\_III.  
 DR InterPro: IPR007861; Muts\_IV.  
 DR Pfam: PF05188; Muts\_II; 1.  
 DR Pfam: PF05192; Muts\_III; 1.  
 DR Pfam: PF05190; Muts\_IV; 1.  
 DR Pfam: PF00488; Muts\_V; 1.  
 DR ProDom: PD001263; Muts\_C; 1.  
 DR SMART: SM00534; Mutsac; 1.  
 DR SMART: SM00533; Mutsd; 1.  
 DR PROSITE: PS00486; DNA\_MISMATCH\_REPAIR\_2; 1.  
 DR Meiosis: ATP-binding; DNA-binding.  
 DR NP\_BIND: 634 641 ATP (POTENTIAL).  
 DR CONFLICT: 811 812 MD -> IH (IN REF. 2 AND 3).  
 FT SEQUENCE 878 AA; 99220 MW; 0FALCIF29CC3DCA CRC64;

Query Match 3.3%; Score 100.5; DB 1; Length 878;  
 Best Local Similarity 19.3%; Pred. No. 16;  
 Matches 87; Conservative 72; Mismatches 182; Indels 109; Gaps 19;

QY 64 DVSTKPNVRYNEN-----FLSLQDPSIIHGMFLSKRLVTLARGISPAFLRF 115  
 DB 115 DISTRIGLCINCTGMYLSDMDSQIYRYVHK--LQYQPTTEILIPSSSLAPTV-- 169  
 QY 116 GGRKTDLPQNLNPAKSGRGGPDYVLKNYEDDIVRSDVAL--DKQCKIAQPDVM 173  
 DB 170 -SKLATMIRK-NVAETVRIKESGRKC--NSQDGLAATIKYIMDTKDKLKEEIIDKT 224  
 QY 174 LVIOREKQAQMHLV-----LIKEQFSNTSNLII--TASLDKLYNFADG 216  
 DB 225 FALCAASAAISYBEIISKSRMNAFRKRIQFEGEENMLIDSKTVAGLEIVEKDKL 284  
 QY 217 SGLHLFALNALRRPNVNSNSSALSLKYSAKKYNTSMEL--GNENNVYRTWNGRA 273  
 DB 285 NGISL-----WKFLDTSTKMGQSLNSITQPLTDGSIEMLEALELK 330  
 QY 274 VNGSQIGKDYIQKSLQPIRIYSRASLYGPNIGRP--RKNVIALLDGEMKVASTVDAV 331  
 DB 331 ANDDLQKRLKEMKSPDLKFLSRLLCINHSAIKPDQRIYVLLIKETLQSVKSLKDAL 390  
 QY 332 TWQHCYIDGIVVVMFLKTRLDL-----SDQIKIQVNTVTPGKKIMEGVVTS 386  
 DB 391 NDQ-----LIQSRLLISTETKIFNNDAIMEIEKLINSCINEDCVASSATQLL 437  
 QY 387 AGGTNNLSDSYAGFLMNTLGMV--ANQGDIVYIRHSF-----DHGYN-- 429  
 DB 438 N-----QRSYA--VXSDNGLLDVSROIYKEVKEEFREVEDLTAQKKINLDHNYDSA 488  
 QY 430 -----HLYDQNFN-----PLPDYMLSLLYKR 450  
 DB 489 RGFYLRIKROEFTDVATLPDVFISRTTK 518

RESULT 14  
 LINE\_HUMAN STANDARD; PRT; 1259 AA.  
 AC P08547;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 01-AUG-1988 (Rel. 08, Last annotation update)  
 DE LINE-1 reverse transcriptase homolog.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;

RP MEDLINE:86230917; PubMed:2423883;  
 RA Hattori M., Kubara S., Takenaka O., Sakaki Y.;  
 RT "L1 family of repetitive DNA sequences in primates may be derived  
 RT from a sequence encoding a reverse transcriptase-related protein";  
 RL Nature 321:625-628(1986).  
 CC -1- MUSELANE005: This sequence was constructed from an alignment of  
 CC published and unpublished sequences, determined in various  
 CC laboratories, belonging to the LINE-1 family.  
 DR PIR: A25313; GNHUL1.  
 DR InterPro: IPR005135; Exo\_endo\_phos.  
 DR InterPro: IPR000477; RYase.  
 DR Pfam: PF03372; Exo\_endo\_phos; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR RNA-directed DNA polymerase.  
 KW RNA-directed DNA polymerase.  
 SQ SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;

Query Match 3.3%; Score 100.5; DB 1; Length 1259;  
 Best Local Similarity 20.0%; Pred. No. 26;  
 Matches 131; Conservative 89; Mismatches 217; Indels 217; Gaps 35;

QY 50 DRAAGLKEKTLIILDVSTKPNVRT-----VNEFLSLQDPSIIHGMFLSKRL 101  
 DB 130 DLORDLDSHTIIMDDFTV--FLSTLDSTRQKINKDQELN--SALHQA-----DL 176  
 QY 102 VTLARGLSPALFRGGRKTDLPQNLNPAKSGRGGPDYVLKNYEDDIVRSDVALDKQ 161  
 DB 177 IDIRTLHP-----KSTETTFP-----SAHHYTSKTDHILGSKTILSKC 216  
 QY 162 KCKCI-----AQHPDMLVIOREKQAQMHLV-LK----- 190  
 DB 217 KRTETITNCSDHSAIKLEIRIKLTQNHSTYTNLNLINDVYHNMKAIEIKFPETN 276  
 QY 191 EGFSTYTNLILTPARSIDKLYNFPADCGGHLIFPLNLLRRPNNSNSSALSLTKXSAS 250  
 DB 277 ENKDTYQNLMDTKAAV-----CRGKFL--ALNHKKRQERS-KITPLISQLK--- 321  
 QY 251 KKYNISWELGNEPNVNTWNGRAVNGSQIGDYIQKSL-----LQPIRIYSBASLYGP 304  
 DB 322 -----ELKQQTGSKASRQEIIRKIRLEKEIETQTLQKIN-ESRSWFPEKI 369  
 QY 305 -NIGRPKNVIA-----LDDGFMKVAS--TVDAVT-----WQCYIDG-RVYVMD 347  
 DB 370 NKIDRPLARLLKKRERKNQIDITINDRGDITTPTEIQTIREYKHLVANKLENEEMD 429  
 QY 348 -FLKRLDLDTSDQ-----IRKIQVNTVTPGKKIMEGVVTSAGGTNNLSDS 386  
 DB 430 KFLDTYLPRLNGEVSINRPITSSSEIHALIN-LPNKK-----SPGEGGFAEF 479  
 QY 397 Y-----AAGFLWNTQGLMANQGDIVYIRHSFD-----HGYNHLYDQNFNPLPDY 442  
 DB 480 YQRYKEBELVPL-LKLPQSIKEKEI--LPNSFEASIIILIPXGRDITTKENFRPIS-- 533  
 QY 443 WLSLLYKRLIPKYLA-----VHVAGLQKRPDGV-----IRDLRIYACCTNNHNH 490  
 DB 534 -LNNIDAKILN-KILANQIQHIIKLIHHDOVGFIPLMGMFNIRKINILQIHNRKOT 591  
 QY 491 NYVGGSTLFTINLHRSRKKIKLAGTLRDLVHQYLLQPYGQEG-----K 536  
 DB 592 NHM-----IISID-----AEKAFDKIQCFPMUKPLNKIGIDTYLAKIRAIYDKP 636  
 QY 537 SKSVQUNGQPLVWDDGTLPELXERPLRAGRTLVIPVTVGFFVYKVNALACR 590  
 DB 637 TANILNQ-----KLEAPLKTGTGTCGCPISPLIPNIVLEVLARAIR 679



Db	2623	SS----	LIFLXPLPGTYQFSLQKSGSDSKIR-----	SACHVAVVIP-----	PTNLTW	2669
Qy	237	NSSSALSLKYASAKKNISWELGNEBNNTYRTMHGRAVN-----	GSOLGXYIQJK			287
Db	2670	NIPSVI----	FAIRRNNT-----	PNLFHLPBGYSLSDDORTFSLISGSGKNSKLS		2717
Qy	288	SLQPIRIYSRASLYGPNIGRPKNVIA--	LIDGFMKVAGSTVDAYTMOHCYIDGRV--			343
Db	2718	SGVYQVNV-----	VGKDEKKEIVRILD-----	DVADDTSKD--	IEYHVSS	2758
Qy	344	-----	KYMDFLKTRLDLTSDQIRKIQK-----	VNTYTPGKKIMLEGVVTTSGG		389
Db	2759	TLSNLKIPFPIDVECPRETEENLEYEITKOCRLFNSDVINTTIP-----	VVTSFANS			2810
Qy	390	TNNL	393			
Db	2811	TNNL	2814			

Search completed: May 6, 2004, 13:44:42  
 Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2004, 13:41:41 ; Search time 45 Seconds  
(without alignments)  
4150.817 Million cell updates/sec

Title: US-10-088-676-2

Perfect score: 3088  
Sequence: 1 MEVLCAFPAMPSSNSRPPA.....PVTMGFYVKNVNALACRFR 592

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MNC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP RODENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_VIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3078	99.7	592	4 Q9HB37	Q9HB37 homo sapien
2	3070	99.4	592	4 Q8WQ2	Q8WQ2 homo sapien
3	2785	90.2	548	4 Q8WQ1	Q8WQ1 homo sapien
4	2736	88.6	534	4 Q9HB38	Q9HB38 homo sapien
5	2446	79.2	480	4 Q9HB39	Q9HB39 homo sapien
6	1655.5	37.7	535	11 Q8XK3	Q8XK3 mus musculu
7	1165	37.7	536	11 Q9QZF8	Q9QZF8 rattus norv
8	1149	37.2	545	4 Q9UD39	Q9UD39 homo sapien
9	1144.5	37.1	543	4 Q9Y251	Q9Y251 homo sapien
10	1138	36.9	545	6 Q9MYX0	Q9MYX0 bos taurus
11	1033.5	33.5	523	13 Q90YX5	Q90YX5 gallus gall
12	625	20.2	515	5 Q8T108	Q8T108 bombyx mori
13	384	12.4	544	10 Q8H615	Q8H615 cryza sativ
14	382	12.4	521	10 Q9SDA1	Q9SDA1 arabidopsis
15	382	12.4	543	10 Q9PFI0	Q9PFI0 arabidopsis
16	334.5	10.8	516	10 Q9FLK8	Q9FLK8 arabidopsis

17	334.5	10.8	539	10 Q8L608	Q8L608 arabidopsis
18	314	10.2	536	10 Q9FZP1	Q9FZP1 arabidopsis
19	307.5	10.0	527	10 Q9RCR8	Q9RCR8 scutellaria
20	301.5	9.8	559	16 Q89F99	Q89F99 bradyrhizob
21	171	5.5	935	5 Q9VZ79	Q9VZ79 drosophila
22	155	5.0	190	10 Q8Z604	Q8Z604 arabidopsis
23	146	4.7	1260	5 Q15639	Q15639 dictyostell
24	128	4.1	1027	16 Q8Z028	Q8Z028 anabaena sp
25	127.5	4.1	1234	5 Q9Y1H7	Q9Y1H7 dictyostell
26	119	3.9	370	17 Q8TP09	Q8TP09 methanosarc
27	118.5	3.8	301	17 Q8TPM8	Q8TPM8 glycine max
28	118.5	3.8	901	10 Q84ZU8	Q84ZU8 glycine max
29	118	3.8	459	10 Q9FVK3	Q9FVK3 glycine max
30	115	3.7	897	10 Q84ZV8	Q84ZV8 glycine max
31	114.5	3.7	3218	5 Q8ILX0	Q8ILX0 plasmodium
32	114	3.7	424	5 Q8T8M7	Q8T8M7 caenorhabdi
33	114	3.7	556	12 Q805J2	Q805J2 strawberry
34	110.5	3.6	1471	5 Q9VU00	Q9VU00 drosophila
35	110	3.6	174	10 Q9ATW5	Q9ATW5 zea mays (m
36	110	3.6	370	17 Q8TR29	Q8TR29 methanosarc
37	110	3.6	370	17 Q8TH26	Q8TH26 methanosarc
38	109	3.5	3317	16 Q8EWP8	Q8EWP8 mycoplasma
39	108.5	3.5	1156	16 Q8KFM2	Q8KFM2 chlorobium
40	107.5	3.5	1053	2 P71329	P71329 fibrobacter
41	107	3.5	895	10 Q84ZY3	Q84ZY3 glycine max
42	107	3.5	915	16 Q7U2L6	Q7U2L6 prochloroco
43	106	3.4	556	12 Q80HR4	Q80HR4 strawberry
44	106	3.4	565	10 Q9LT48	Q9LT48 arabidopsis
45	106	3.4	762	5 Q27701	Q27701 macrobdeilia

## ALIGNMENTS

RESULT 1  
ID Q9HB37 PRELIMINARY: PRT; 592 AA.

AC Q9HB37; 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Hepatanase-like protein HPA2C.  
DR Hepatanase-1 like protein HPA2C.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20483645; PubMed=11027606;  
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,  
RA Hironaka M., Patel S., Barry R., Stuberfield C., Terrett J., Page M.,  
RT "Cloning and Expression Profiling of Hpa2, a Novel Mammalian  
RT Hepatanase Family Member."  
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).  
DR EMBL; AF282887; AAG2423.1; --  
DR GO; GO:0005622; C:intracellular; TAS.  
DR GO; GO:0030305; F:hepatanase activity; TAS.  
DR InterPro; IPR005199; Glyco\_hydro\_79n.1.  
DR Pfam; PF03662; Glyco\_hydro\_79n.1.  
SQ SEQUENCE 592 AA; 66580 MW; 95C384AD9A74258E CRC64;

Query Match 99.7%; Score 3078; DB 4; Length 592;  
Best Local Similarity 99.7%; Pred. No. 1.8e-235;  
Matches 590; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEVLCAFPAMPSSNSRPPACIAPGALYALILHLISQAGDRRLPVDBAAGKENTL 60  
DB 1 MEVLCAFPAMPSSNSRPPACIAPGALYALILHLISQAGDRRLPVDBAAGKENTL 60  
QY 61 ILLDVSTKNPVATVENFSLQIDPSIHDGMDLFLSKRLVTTARGSPAPLRGGRKT 120  
DB 61 ILLDVSTKNPVATVENFSLQIDPSIHDGMDLFLSKRLVTTARGSPAPLRGGRKT 120

```

QY 121 DFLQFQNLNRPAAKSRGGPGPDYLLKNYEDDIVRSVDVALDKXGCKIAQHPDVMVLQREK 180
DB 121 DFLQFQNLNRPAAKSRGGPGPDYLLKNYEDDIVRSVDVALDKXGCKIAQHPDVMVLQREK 180
QY 181 AAOQMLVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIPALNALRRPNNSWNSS 240
DB 181 AAOQMLVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIPALNALRRPNNSWNSS 240
QY 241 ALSLKYSASKKNISWELGNEBNRYRWGRAVNSQGLKDYIQLKSLLOPIRISRAS 300
DB 241 ALSLKYSASKKNISWELGNEBNRYRWGRAVNSQGLKDYIQLKSLLOPIRISRAS 300
QY 301 LYGPNIRPRKNVIALDGMKVAAGSTVDVAVTMOHCYIDGRVYKWDPLKTRLLDITLSDQ 360
DB 301 LYGPNIRPRKNVIALDGMKVAAGSTVDVAVTMOHCYIDGRVYKWDPLKTRLLDITLSDQ 360
QY 361 IRKIÖKVNTYTPGKKIWEGVVTSAGGTTNLSDSYAAGFLMLNTLGMLANOGIDVYIR 420
DB 361 IRKIÖKVNTYTPGKKIWEGVVTSAGGTTNLSDSYAAGFLMLNTLGMLANOGIDVYIR 420
QY 421 HSPFDHGYNHLVDQNFNPDPYMLSLLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
DB 421 HSPFDHGYNHLVDQNFNPDPYMLSLLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
QY 481 YAHCTNNHNNHYVRSITLFTINLHRSRKKIKLAGTIRDKLVHGYLLQPYGQGLSKSV 540
DB 481 YAHCTNNHNNHYVRSITLFTINLHRSRKKIKLAGTIRDKLVHGYLLQPYGQGLSKSV 540
QY 541 QLNQGPLVWDDGTLPELKRPRLRAGRTLYIPPTMGFFVYKVNALACRYR 592
DB 541 QLNQGPLVWDDGTLPELKRPRLRAGRTLYIPPTMGFFVYKVNALACRYR 592

```

## RESULT 2

```

ID Q8MWQ2 PRELIMINARY; PRT; 592 AA.
AC Q8MWQ2;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Heparanase 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Legoux P., Legoux R., O'Brien D., Salome M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299719; CAC82491.1;
DR Genev; HGNC:18374; HPSR2.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 592 AA; 66520 MW; 9478841FEACD558B CRC64;

Query Match 99.4%; Score 3070; DB 4; Length 592;
Best Local Similarity 99.4%; Pred. No. 7.9e-235;
Matches 599; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 121 DFLQFQNLNRPAAKSRGGPGPDYLLKNYEDDIVRSVDVALDKXGCKIAQHPDVMVLQREK 180
DB 121 DFLQFQNLNRPAAKSRGGPGPDYLLKNYEDDIVRSVDVALDKXGCKIAQHPDVMVLQREK 180
QY 181 AAOQMLVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIPALNALRRPNNSWNSS 240
DB 181 AAOQMLVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIPALNALRRPNNSWNSS 240
QY 241 ALSLKYSASKKNISWELGNEBNRYRWGRAVNSQGLKDYIQLKSLLOPIRISRAS 300
DB 241 ALSLKYSASKKNISWELGNEBNRYRWGRAVNSQGLKDYIQLKSLLOPIRISRAS 300
QY 301 LYGPNIRPRKNVIALDGMKVAAGSTVDVAVTMOHCYIDGRVYKWDPLKTRLLDITLSDQ 360
DB 301 LYGPNIRPRKNVIALDGMKVAAGSTVDVAVTMOHCYIDGRVYKWDPLKTRLLDITLSDQ 360
QY 361 IRKIÖKVNTYTPGKKIWEGVVTSAGGTTNLSDSYAAGFLMLNTLGMLANOGIDVYIR 420
DB 361 IRKIÖKVNTYTPGKKIWEGVVTSAGGTTNLSDSYAAGFLMLNTLGMLANOGIDVYIR 420
QY 421 HSPFDHGYNHLVDQNFNPDPYMLSLLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
DB 421 HSPFDHGYNHLVDQNFNPDPYMLSLLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
QY 481 YAHCTNNHNNHYVRSITLFTINLHRSRKKIKLAGTIRDKLVHGYLLQPYGQGLSKSV 540
DB 481 YAHCTNNHNNHYVRSITLFTINLHRSRKKIKLAGTIRDKLVHGYLLQPYGQGLSKSV 540
QY 541 QLNQGPLVWDDGTLPELKRPRLRAGRTLYIPPTMGFFVYKVNALACRYR 592
DB 541 QLNQGPLVWDDGTLPELKRPRLRAGRTLYIPPTMGFFVYKVNALACRYR 592

```

## RESULT 3

```

ID Q8MWQ1 PRELIMINARY; PRT; 548 AA.
AC Q8MWQ1;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 24, Last annotation update)
DE Heparanase 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Legoux P., Legoux R., O'Brien D., Salome M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299720; CAC82492.1;
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 548 AA; 61771 MW; B8986303RC73A60A CRC64;

Query Match 90.2%; Score 2785; DB 4; Length 548;
Best Local Similarity 99.3%; Pred. No. 2.9e-212;
Matches 535; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Db 121 DFLQFONLRNPAKSRGGPGPDYLLKXVEDDIVASDVALDKQCKIAQHPDVLLEIQREK 180
QY 181 AAOQMHVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIFALNALRRPNNSNSS 240
Db 181 AAOQMHVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIFALNALRRPNNSNSS 240
QY 241 ALSLKYSASKKYNISWELGNEPNRYTMGRAVNSQSGKDYIOLKSLQPIRISRAS 300
Db 241 ALSLKYSASKKYNISWELGNEPNRYTMGRAVNSQSGKDYIOLKSLQPIRISRAS 300
QY 301 LYPENIGRPRKNVIALLDGFMKVAAGSTVDATWQHCHIDGRVYKVMDFLKTLLDITLSDQ 360
Db 301 LYPENIGRPRKNVIALLDGFMKVAAGSTVDATWQHCHIDGRVYKVMDFLKTLLDITLSDQ 360
QY 361 IRKIQKVNNTYTPGKIMLEGVVTTSGAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 420
Db 361 IRKIQKVNNTYTPGKIMLEGVVTTSGAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 420
QY 421 HSFFDHGYNHLDVONFNPDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
Db 421 HSFFDHGYNHLDVONFNPDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
QY 481 YACHTNHNHYVRSITLFTIINHRSRKKIKLAGTRDKLVHGYLLQPYQSGELKSKS 539
Db 481 YACHTNHNHYVRSITLFTIINHRSRKKIKLAGTRDKLVHGYLLQPYQSGELKSKT 539

```

## RESULT 4

```

QY 09HB38 PRELIMINARY; PRT; 534 AA.
AC 09HB38;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hepatanase-like protein HPA2p.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20483645; PubMed=11027606;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hircock M., Patel S., Barry E., Stubberfield C., Terrett J., Page M.;
RT "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT Hepatanase Family Member.";
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR EMBL; AF282886; AAG23422.1;
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 534 AA; 60063 MW; C3DE5E900CB38C4 CRC64;

```

Query Match 88.6%; Score 2736; DB 4; Length 534;  
 Best Local Similarity 89.9%; Pred. No. 2.2e-208; Indels 58; Gaps 1;  
 Matches 532; Conservative 1; Mismatches 1;

```

QY 1 MRVLCAPPEAMPSSNSRPACLAAGALYALLLHLSSQAGDRRLPVDRAGLKEKTL 60
Db 1 MRVLCAPPEAMPSSNSRPACLAAGALYALLLHLSSQAGDRRLPVDRAGLKEKTL 60
QY 61 ILDDVSTKNPRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFRLFQGKRT 120
Db 61 ILDDVSTKNPRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFRLFQGKRT 120
QY 121 DFLQFONLRNPAKSRGGPGPDYLLKXVEDDIVASDVALDKQCKIAQHPDVLLEIQREK 180
Db 121 DFLQFONLRNPAKSRGGPGPDYLLKXVEDDIVASDVALDKQCKIAQHPDVLLEIQREK 180
QY 181 AAOQMHVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIFALNALRRPNNSNSS 240
Db 181 AAOQMHVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIFALNALRRPNNSNSS 240

```

```

QY 241 ALSLKYSASKKYNISWELGNEPNRYTMGRAVNSQSGKDYIOLKSLQPIRISRAS 300
Db 241 -----EPNNRYTMGRAVNSQSGKDYIOLKSLQPIRISRAS 242
QY 301 LYPENIGRPRKNVIALLDGFMKVAAGSTVDATWQHCHIDGRVYKVMDFLKTLLDITLSDQ 360
Db 301 LYPENIGRPRKNVIALLDGFMKVAAGSTVDATWQHCHIDGRVYKVMDFLKTLLDITLSDQ 360
QY 361 IRKIQKVNNTYTPGKIMLEGVVTTSGAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 420
Db 361 IRKIQKVNNTYTPGKIMLEGVVTTSGAGTNNLSDSAAGFLMNTLGMLANOGIDVIR 420
QY 421 HSFFDHGYNHLDVONFNPDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
Db 421 HSFFDHGYNHLDVONFNPDPYMLSLYKRLIGPKYLAHVAGLQKRPGRVIRDKLRI 480
QY 481 YACHTNHNHYVRSITLFTIINHRSRKKIKLAGTRDKLVHGYLLQPYQSGELKSKS 539
Db 481 YACHTNHNHYVRSITLFTIINHRSRKKIKLAGTRDKLVHGYLLQPYQSGELKSKS 539

```

## RESULT 5

```

QY 09HB39 PRELIMINARY; PRT; 480 AA.
AC 09HB39;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hepatanase-like protein HPA2a.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20483645; PubMed=11027606;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hircock M., Patel S., Barry E., Stubberfield C., Terrett J., Page M.;
RT "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT Hepatanase Family Member.";
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR EMBL; JG7506; JG7506.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 480 AA; 53900 MW; F75F89F67AC1FF63 CRC64;

```

Query Match 79.2%; Score 2446; DB 4; Length 480;  
 Best Local Similarity 80.9%; Pred. No. 1.9e-185; Indels 112; Gaps 1;  
 Matches 479; Conservative 1; Mismatches 0;

```

QY 1 MRVLCAPPEAMPSSNSRPACLAAGALYALLLHLSSQAGDRRLPVDRAGLKEKTL 60
Db 1 MRVLCAPPEAMPSSNSRPACLAAGALYALLLHLSSQAGDRRLPVDRAGLKEKTL 60
QY 61 ILDDVSTKNPRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFRLFQGKRT 120
Db 61 ILDDVSTKNPRTVNTENFLSLQDPSIIHGMDFLSSKRLVTLARGLSPAFRLFQGKRT 120
QY 121 DFLQFONLRNPAKSRGGPGPDYLLKXVEDDIVASDVALDKQCKIAQHPDVLLEIQREK 180
Db 121 DFLQFONLRNPAKSRGGPGPDYLLKXVEDDIVASDVALDKQCKIAQHPDVLLEIQREK 180
QY 181 AAOQMHVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIFALNALRRPNNSNSS 240
Db 181 AAOQMHVLLKEQFSNTYSNLTARSLDKYNFADCSGHLIFALNALRRPNNSNSS 240

```

```

Db 150 -----EENNRYTMHGRAVNSQLGKDYIQLKSLIQIRIYSRA 188
QY 301 LYGPNIGRPRKNVIALLDGFMKVAAGSTVDVAVTMOHCYIDGRVYKMDPLKTLDTLSDQ 360
Db 189 LYGPNIGRPRKNVIALLDGFMKVAAGSTVDVAVTMOHCYIDGRVYKMDPLKTLDTLSDQ 248
QY 361 IRIQVNVNTYTPGKKIMLEGVVTTTSAAGTNNLSDSYAGFLMNTLGLANQSIDVYIR 420
Db 249 IRIQVNVNTYTPGKKIMLEGVVTTTSAAGTNNLSDSYAGFLMNTLGLANQSIDVYIR 308
QY 421 HSPFDGYNHLYVDQNFNPLPDYWLSLYKRLIGPKYLAHVAGLQKPRPGVIRDKLRI 480
Db 309 HSPFDGYNHLYVDQNFNPLPDYWLSLYKRLIGPKYLAHVAGLQKPRPGVIRDKLRI 368
QY 481 YACHTNHNHNYRGSITLFTIIMHRSKKIKLAGTLRDKLVHQYLLQPYGQGLKSKSV 540
Db 369 YACHTNHNHNYRGSITLFTIIMHRSKKIKLAGTLRDKLVHQYLLQPYGQGLKSKSV 428
QY 541 QLNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592
Db 429 QLNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 480

```

## RESULT 6

```

Q8K3K3 PRELIMINARY; PRT; 535 AA.
ID Q8K3K3
AC Q8K3K3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Heparanase.
GN HPSE OR HPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxId=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=FEB;
RA Mao H.-Q., Navarro E., Patel S., Sargent D., Koo H., Wan H.,
RA Plata A., Zhou Q., Ludwig D., Bohlen P., Kusie P.,
RT "Cloning, expression, and purification of mouse heparanase.";
RL Protein Expr. Purif. 0:0-0(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AY07467; AAL76083.1; -.
DR EMBL; AK040471; BAC30600.1; -.
DR MGD; MGI:1343124; Hpsa.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 535 AA; 60065 MW; 6E73A8302FB8A0DF CRC64;

```

```

Query Match 37.7%; Score 1165.5; DB 11; Length 535;
Best Local Similarity 46.2%; Pred. No. 9.9e-84;
Matches 244; Conservative 81; Mismatches 176; Indels 27; Gaps 6;
QY 63 LDVSTNPTVVENFSLQDPSIITH-GWLDPLSKRLVTLARGSLPAFRFGKRTD 121
Db 33 LEVYTRPARSVPSTLITIDSLATDPRFLTFLGSPRLARSLRAYIRFGGRTD 92
QY 122 FLOFQNLNPAKSGRGPDPYILKNEYDIIVRSVDALDKQKCKIAQHPDVLVIQREKA 181
Db 93 FLIF-----DPDKPTSEERSYWSQVNHDIQSEEPV-----SAAVLRKLQVEWP 137
QY 182 AQCHLVILKEQFNTSNILITARSIDKLYNFPADCGSLIFALNALRNPNNSWSSA 241

```

```

Db 138 FQ-ELLIREQYQKEFNSTYSRSSVDMLXSPAKCGDLIRGLNALLRTPDLRMSSNA 196
QY 242 LSLIKYSASKKYNISWELGNEPNNRYTMHGRAVNSQLGKDYIQLKSLQIRIYSRA 301
Db 197 QLLIDYSSSKYINISWELGNEPNSFWKKAIIIDGLQGLGDEVELHKLQK-SAFONAKI 255
QY 302 YGNPGRPRKNVIALLDGFMKVAAGSTVDVAVTMOHCYIDGRVYKMDPLKTLDTLSDQ 361
Db 256 YGNPGRPRKNVIALLDGFMKVAAGSTVDVAVTMOHCYIDGRVYKMDPLKTLDTLSDQ 315
QY 362 RIKQVNVNTYTPGKKIMLEGVVTTTSAAGTNNLSDSYAGFLMNTLGLANQSIDVYIR 421
Db 316 RIKQVNVNTYTPGKKIMLEGVVTTTSAAGTNNLSDSYAGFLMNTLGLANQSIDVYIR 375
QY 422 SFPDGHNLVYDQNFNPLPDYWLSLYKRLIGPKYLAHVAGLQKPRPGVIRDKLRI 481
Db 376 SFPDGHNLVYDQNFNPLPDYWLSLYKRLIGPKYLAHVAGLQKPRPGVIRDKLRI 426
QY 482 AHCNTNHNHNYRGSITLFTIIMHRSKKIKLAGTLRDKLVHQYLLQPYGQGLKSKSV 541
Db 427 AHCNTNHNHNYRGSITLFTIIMHRSKKIKLAGTLRDKLVHQYLLQPYGQGLKSKSV 486
QY 542 LNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 592
Db 487 LNGQPLVWDDGTLPELKRPLRAGRTLVIPVTMGFYVYKVNALACRYR 534

```

## RESULT 7

```

Q9QZF8 PRELIMINARY; PRT; 536 AA.
ID Q9QZF8
AC Q9QZF8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Heparanase.
GN HEP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxId=10116;
RN [1]
RN SEQUENCE FROM N.A.
RA Podyma K.A., Yokote H., Sakaguchi K., Ikura M., Yanagishita M.;
RA "Heparanase from parathyroid cell line";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184967; AAF04563.1; -.
DR InterPro; IPR005199; Glyco_hydro_79N.
DR Pfam; PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 536 AA; 60569 MW; 6208B1FD9E28421 CRC64;

```

```

Query Match 37.7%; Score 1165; DB 11; Length 536;
Best Local Similarity 45.1%; Pred. No. 1.1e-83;
Matches 248; Conservative 87; Mismatches 187; Indels 28; Gaps 7;
QY 42 GDRRLPLVDPAAG-LKEKTLILDVSTNPTVVENFSLQDPSIITH-GWLDPLSKR 99
Db 12 GRNALNQGPRPAGAPKVDVVDLEFTRKRLFGQVSPSTITDASLATDPRFLTFLGSP 71
QY 100 RLVTIARGSLPAFRFGKRTDPLQFQNLNPAKSGRGPDPYILKNEYDIIVRSVDALD 159
Db 72 RLRLARGSLPAFRFGKRTDPLIF-----DPDKPTSEERSYWSQVNHDIQSEEPV----- 123
QY 160 KQKCKIAQHPDVLVIQREKAQMHLYLKEQFNTYSRLITARSIDKLYNFPADCGSL 219
Db 124 -----RVSADVLRKLQVEWPQ-ELLILREQYQKEFNSTYSRSSVDMLXSPAKCGDLIRGLNALLRTPDLRMSSNA 175
QY 220 HLITNALNARNPNNSWSSALSILKYSASKKYNISWELGNEPNNRYTMHGRAVNSQL 279
Db 176 DLIFGNALNRTDPLRNSSNAQLILNLYCSSKYNISWELGNEPNSFWKKAQISIDGLQ 235
QY 280 GKDYIQLKSLQIRIYSRAISLIGPNIGRPRKNVIALLDGFMKVAAGSTVDVAVTMOHCYID 339
Db 236 GEDVELHKLQK-SAQNKAKLIGPDIQGPRTGTVLIRSFLLAGGVIDSLTWHHYLYN 294

```



QY 340 GRVVKWDFKTRLLDLSQIRKIQKVVNTYTPGKKIWLGVVTSAGTNNLSDSYAA 339  
 DB 295 GRVATKEDPSSDVLDFIISVQKILKATKEMTPGKVMIGETSAAVGAGAPLISNTEFA 354  
 QY 400 GFLMLNTLGMLANOGIDIVYHRSFDFHGNHLVDQNNPLPDVYLSLLYKRLGPKVLAV 459  
 DB 355 GFWMLDLKGLSAQIGIEVVMQVFFGAGNVHLVDENFEPDPVYLSLFFKVLGPKVLM 414  
 QY 460 HVAGLQKRPGRGVYRDKLRIYACTNNHNNVYRGSITLLFIINLHRSKKIKLAGTLRD 519  
 DB 415 RVKRPD-----RSKLRVYLAHCTNHYHRYREGDLTYLVNLHVHTLKLPPMPFS 465  
 QY 520 KLVHGYLLQPYGQGLSKSVQVNGQPLVWVDGTLBELKRPRLPAGRTVLPVTMGFY 579  
 DB 466 RPVKRYLLKPGSGGLSKSVQVNGQTLKVDQTLPALTEKPLPAGSSISVPAFSYGF 525  
 QY 580 VVKVNNALAC 589  
 DB 526 VIRNAKTIAC 535

## RESULT 8

Q9UL39 PRELIMINARY; PRT; 545 AA.

AC 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Heparanase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=2029546; PubMed=10764835;  
 RA Dempsey L.A., Plummer T.B., Coombes S.L., Platt J.L.;  
 RT "Heparanase expression in invasive trophoblasts and acute vascular  
 damage.";  
 RL Glycobiology 10:467-475 (2000).  
 DR EMBL; AF084467; AAS4516.1; -;  
 DR InterPro; IPR005199; Glyco\_hydro\_79n.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 SQ SEQUENCE 545 AA; 61418 MW; 67B80ACD73C5A9A1 CRC64;

Query Match 37.2%; Score 1149; DB 4; Length 545;

Best Local Similarity 43.7%; Pred. No. 2, 1e-82;  
 Matches 251; Conservative 82; Mismatches 206; Indels 36; Gaps 8;

QY 17 RPPACIACGATYLLALHLSSSQAGDRRPLPVDRAAGLKEKTLILDVSTKNPRTVNE 76  
 DB 4 RSKRPLPPLMLLILGLPLSPALRPAQA-----QDVVDLDFPQBLHLVSP 56  
 QY 77 NFLSLQDLPSTIHD-GWIDFLSSKRLVTLARGLSAFIRFGKRTDPLQFQMLNPKSR 135  
 DB 57 SFLSTYITANLATDRFLILGLSPKRLTARGLSPAYIRFGTKTDFILF---DPKES 112  
 QY 136 GGPQGDVYLYKVEDDIVSDVALDKQKCKIAQ-HPDVMLYLOEKAAQMLVLLKEQFS 194  
 DB 113 TPEERSVQSQVNOI-----CKGSIPTDVEEKLRLPEWYQGL-LREHYQ 159  
 QY 195 NTYSWLLITANSLLDYLPADCSGLHLIFALNALRRNNNSWSSALSILKYASAKYN 254  
 DB 160 KKFKNSTYRSRSVDVLYTFANCOSGLDLIFGLNALRLTDLQWNSNAQLLDYCSKGYN 219  
 QY 255 ISMELGNENPNRTWAGRVANGSOLGKYIOLKSLQPIRTYSRASTVGNIGRPKKVI 314  
 DB 220 ISMELGNENPNRTWAGRVANGSOLGKYIOLKSLQPIRTYSRASTVGNIGRPKKVI 278  
 QY 315 ALLDGFMYAGSTVDAVTVQHCYIDGRVVKWDFKTRLLDLSQIRKIQKVVNTYTPG 374

DB 279 KMLKSFLLKAGEVIDSVTHHHYVINGRTATREDFLNPVLDIFISSVQKQVVESTIRPG 338  
 QY 375 KXIWLEGVVTSAGTNNLSDSYAGFLMLNTLGMLANOGIDIVYHRSFDFHGNHLVDQ 434  
 DB 339 KKVWLGERTSAAVGAGAPLISDTPAAGFWMLDKLGLSANGIEVVMQVFFGAGNVHLVDE 398  
 QY 435 NNPNLPDVTWLSLLYKRLGPKVLAVHAGLQKRPGRGVYRDKLRIYACTNNHNNVYR 494  
 DB 399 NFDPLPDVYLSLFFKVLGPKVLMASVQSSKR-----KLRVYLAHCTNIDPRYKE 449  
 QY 495 GSITLLFIINLHRSKKIKLAGTLRDKLVHGYLLQPYGQGLSKSVQVNGQPLVWVDGT 554  
 DB 450 GDLITLAIKLVHVTYLRPLYPFSNKQYDKILRPLGPHGLSKSVQVNGTLTKVVDQT 509  
 QY 555 LPBLKRPRLPAGRTVLPVTMGFYVVKVNNALAC 589  
 DB 510 LPPLMEKPLRPGSSILGPAFSYFVIRNAKVIAC 544

## RESULT 9

Q9Y251 PRELIMINARY; PRT; 543 AA.

AC 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE HEPARANASE (HPESE protein).  
 GN HPA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=99321249; PubMed=10395326;  
 RA Hulett M.D., Freeman C., Handorf B.J., Baker R.T., Harris M.J.,  
 RA Parish C.R.;  
 RT "Cloning of mammalian heparanase, an important enzyme in tumor  
 invasion and metastasis.";  
 RL Nat. Med. 5:803-809 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Vlodavsky I., Friedman Y., Elkin M., Aingorn H., Azmon R.,  
 RA Izhai-Michaeli R., Bitan M., Pappo O., Peretz T., Michael I.,  
 RA Spector L., Becker I.;  
 RT "Mammalian heparanase: a novel gene involved in tumor progression and  
 metastasis.";  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99377052; PubMed=10446189;  
 RA Toyoshima M., Nakajima M.;  
 RT "Human heparanase. Purification, characterization, cloning, and  
 expression.";  
 RL J. Biol. Chem. 274:24153-24160 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=9935379; PubMed=10405343;  
 RA Kusiste P.H., Hulmes J.D., Ludwig D., Patel S., Navarro E.C.,  
 RA Seddon A.P., Giorgio N.A., Bohlen P.;  
 RT "Cloning and Functional Expression of a Human Heparanase Gene.";  
 RL Biochem. Biophys. Res. Commun. 261:183-187 (1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Collins B.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh P.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinoli P., Prange C.,  
 RA Raba S.S., Loggellano N.A., Peters G.J., Abrahamson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumatte P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Mizny D.M., Sodergren E.U., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez M.I., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 RA Jones S.J., Maier W.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [6]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;

RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF165154; AAD45379.1; -

DR EMBL; AF144325; AAD4342.1; -

DR EMBL; AF155510; AAD54941.1; -

DR EMBL; AF152376; AAD4669.1; -

DR EMBL; BC051321; AAH51321.1; -

DR Gene; HGNC:5164; HPSB.

DR GO; GO:0004566; F-beta-glucuronidase activity; TAS.

DR GO; GO:0007125; P:invasive growth; TAS.

DR InterPro; IPR005199; Glyco\_hydro\_79N.

DR Pfam; PF03662; Glyco\_hydro\_79n; I.

SQ SEQUENCE 543 AA; 61176 MW; AD262EC267334AB2 CRC64;

Query Match 37.1%; Score 1144.5; DB 4; Length 543;  
 Best Local Similarity 43.2%; Pred. No. 4,7e-82;  
 Matches 248; Conservative 83; Mismatches 190; Indels 53; Gaps 9;

QY 18 PACACGATLALHLHSLSSQAGRRPLPYDRAAGLKEKTLILDVSTKNPRTVEN 77  
 DB 20 PLGSPFGL-----PRPA-----QAGVVDLDFTFQFLHSPS 55  
 QY 78 FLTSLDPSIHD-GWLDPLSKRLVTLARGLSPALFPGKRTDPLQFONLRNPAKRG 136  
 DB 56 FLSVITDALATDPFLLILGSPKRLTLARGLSPALFPGKRTDPLIF---DKKEST 111  
 QY 137 GGPDPYLLKNVEDDVRSDVALDKQKCKIAQ-HEDVWLVLQREKAAQMHVLVLEKFSN 195  
 DB 112 FEERSYWGQVNDI-----CKYGSIPDVEEKLLEPFQEQD-LLEHYQK 158  
 QY 196 TYSNLTILTRSLDKLYNPADCSGLHIFALNALRPNPNNSNSALSLKYSASKYNI 255  
 DB 159 KPNSTYSRSDVLYTFANCSGLDIFELNALTLTADLQNNSSNAQLLDYCSKGYNI 218  
 QY 256 SWELGNEPNNTYTMGRAVNSQLGKDYIQLKSLIQPIRIYRASLYGPNIGRPKXIA 315  
 DB 219 SWELGNEPNNSFKKADIFINSQLEDFTQLHKLIRK-STFNAAKLYGPDVQGPFRKPAK 277  
 QY 316 LLDGFKVAGSTVDATVWQHCYIDGRVYVVMDFLTRLIDLTLSDQIRIKQKVVNYTPGK 375  
 DB 278 MLKSLFKAGGEVIDSTVTHHYLLNGTARFEDFLMPDVLDFISSVQKQFVVESTREK 337  
 QY 376 KWLBSVVTTSAGGTNNLSDSYAAGFLMNTLGMLANQIDVYIHSFFDGHYHLYVN 435  
 DB 338 KWLBSVVTTSAGGTNNLSDSYAAGFLMNTLGMLANQIDVYIHSFFDGHYHLYVN 397  
 QY 436 FNLDPYMLSLYKRLIGPKVLAHVAGLQRPGRVIRDLRLIYAHCTNNHNHNYVG 495  
 DB 398 FNLDPYMLSLYKRLIGPKVLAHVAGLQRPGRVIRDLRLIYAHCTNNHNHNYVG 448  
 QY 496 STTLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYGOBGLKSKSVQNLQPLVMYDDGL 555  
 DB 449 DTLVAIINLHNTYKRLIPYPSNKKQVDKYLRLRPGHLLSKSVQNLQPLVMYDDGL 508

QY 556 PELKPRPLRAGRTLVIPYTMGFWYVKNVNALAC 589  
 DB 509 PFLMEKPLRPGSSILGPASYSFVIRNKNVAAAC 542

RESULT 10

Q9MYT0

ID Q9MYT0

AC Q9MYT0

DT 01-OCT-2000 (T-EMBLrel. 15, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)

DE Hepatanae.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;

OC Bovidae; Bovinae; Bos.

NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Kizaki K., Nakano H., Takahashi T., Imai K., Hashizume K.;

RT "Expression of Hepatanae mRNA in Bovine Placenta During Gestation.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF281160; AAF87301.2; -

DR InterPro; IPR005199; Glyco\_hydro\_79N.

DR Pfam; PF03662; Glyco\_hydro\_79n; I.

SQ SEQUENCE 545 AA; 61076 MW; FAC4BDFD855B933 CRC64;

Query Match 36.9%; Score 1138; DB 6; Length 545;  
 Best Local Similarity 42.9%; Pred. No. 1.6e-81;  
 Matches 248; Conservative 85; Mismatches 201; Indels 44; Gaps 10;

QY 20 ACIAPGATLALHLHSLSSQAGRRP-----LPYDRAAGLKEKTLILDVSTKNPRT 73  
 DB 3 ACKRPG-LRPEPLLLPLGLPLCPSPGPAAAPADDAE-----LEFFTERPLHL 53  
 QY 74 VNEFSLDPSIHD-GWLDPLSKRLVTLARGLSPALFPGKRTDPLQFONLRNPA 132  
 DB 54 VSPAFSLFTDALATDPFLLILGSPKRLTLARGLSPALFPGKRTDPLIFDPKXEA 113  
 QY 133 KSGGPGDPYLLKNVEDDVRSDVALDKQKCKIAQHP-DVWLVLQREKAAQMHVLVLEK 191  
 DB 114 FEB-----RSYMLQSQVNDI-----CKYGSIPDVEEKLLEPFQEQD-LLEHY 156  
 QY 192 QFSNLTILTRSLDKLYNPADCSGLHIFALNALRPNPNNSNSALSLKYSASKYNI 251  
 DB 157 QIOKFTNSTRSDVLYTFANCSGLDIFELNALTLTADLQNNSSNAQLLDYCSK 216  
 QY 252 KYNISWELGNEPNNTYTMGRAVNSQLGKDYIQLKSLIQPIRIYRASLYGPNIGRPK 311  
 DB 217 NYNISWELGNEPNNSFKKADIFINSQLEDFTQLHKLIRK-SAFNAAKLYGPDVQGP 275  
 QY 312 NVIALLDGFKVAGSTVDATVWQHCYIDGRVYVVMDFLTRLIDLTLSDQIRIKQKVV 371  
 DB 276 NYVAMLSFKAGGEVIDSTVTHHYLLNGTARFEDFLMPDVLDFISSVQKRLIVEKI 335  
 QY 372 TPGRKIVLBSVVTTSAGGTNNLSDSYAAGFLMNTLGMLANQIDVYIHSFFDGHYH 431  
 DB 336 RPLKVVLLGTSAGGAFGLFNTFAAGFMWLDKGLSARMGIEVVMQVLFAGANYHL 395  
 QY 432 VDQNFNLPRYMLSLYKRLIGPKVLAHVAGLQRPGRVIRDLRLIYAHCTNNHNH 491  
 DB 396 VDQNFNLPRYMLSLYKRLIGPKVLAHVAGLQRPGRVIRDLRLIYAHCTNNHNH 446  
 QY 492 YVRSITLFTINLHRSRKKIKLAGTRDKLVHGYLLQPYGOBGLKSKSVQNLQPLVMY 551  
 DB 447 YKESDILVYALNHLNHTYKRLIPYPSNKKQVDKYLRLRPGHLLSKSVQNLQPLVM 506  
 QY 552 DGTPELKPRPLRAGRTLVIPYTMGFWYVKNVNALAC 589  
 DB 507 EQTLPALTEKPLRPGSSILGPASYSFVIRNKNVAAAC 544

```

RESULT 11
ID 090YKS PRELIMINARY; PRT; 523 AA.
AC 090YKS;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Heparanase.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11387326;
RA Goldsmith O., Zcharia E., Aingorn H., Guatta-Rangini Z., Atzmon R.,
RA Michal I., Becker I., Mitrani E., Vlodavsky I.,
RT "Expression Pattern and Secretion of Human and Chicken Heparanase Are
RT Determined by Their Signal Peptide Sequence."
RL J. Biol. Chem. 276:29178-29187(2001).
DR EMBL: AY037807; AA082548.1;
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 523 AA; 58386 MW; 8EB0B7B18C9BF881 CRC64;

Query Match 33.5%; Score 1033.5; DB 13; Length 523;
Best Local Similarity 40.8%; Pred. No. 2,9e-73;
Matches 233; Conservative 82; Mismatches 199; Indels 57; Gaps 11;

QY 27 LVALILHLSLSQAGRRPLPYDRAAGLKEKTLILLDSTKNPVTNNENFSLQDLS 86
DB 1 MVLVLLLVLLA-----VPPRRTABLQ-----GLRPFICAVSPAFSLTLDAS 44
QY 87 IINH-CMLDFLSKRLVTLARGLSPAFRFGKRTDFLOFONLRNPAKRGSGPDIYLK 145
DB 45 LADPFPVALLRPKLTHTLASGLSPGLRFGSTPFLF---NPKD-----S 90
QY 146 NYEDIVRSVALDKQKCK-----IAQHPDVLV---LQREKAQMLVLLKEQFSNTYS 198
DB 91 TWEKTLSEFOAKDV---CEAWPSFAVVPKLLITQWPLQEK-----LLLAHSHWKKK 140
QY 199 NILTARSIDLKYNPADCGSLHLIFALNLRPNPNSSSSALLSLKXSASAKKNISWE 258
DB 141 NNTTSTSLDILHTFPASSSGFLVFGINALLRPAQLQWSSNAKOLLGYCAQSYNISWE 200
QY 259 LGNEPNRYRTMGRAVNGSOLGKDYIQLKSLQPIRIYSRASLYGPNISPRKNVITALLD 318
DB 201 LGNEPNSFRKSGICIDGFLGRDPVHLRQLSQHPLRYRAELVGLDVQGRKHTQHLLR 260
QY 319 GPNKVASGTDAVWQHCYIDGRVVKMDFLTKRLLDLSDQIKQKQVNTYTPGKTIW 378
DB 261 SFKSGGKALDSVTMHHYVNGSATREBPLSDPEVLDSPATALHDVLGIVEATVPKQKW 320
QY 379 LBEVVTTSAGCTNNLSDSYAAGFLWNTLQMLNOCGIDVIVIRASFPHGYNHLVDQNFNP 438
DB 321 LGSTGASVGGAPQLSTVYAGFMWMDKGLAARGLDVMRVQSGASHYLVDSGFNP 380
QY 439 LPTVWLSLYKRLIGPVTAAVAVAGLQKRPGRVIRDKRLIAAGCTNNHNNHYVAGSTT 498
DB 381 LPTVWLSLYKRLIGPVTAAVAVAGLQKRPGRVIRDKRLIAAGCTNNHNNHYVAGSTT 431
QY 499 LFTINHSRKTKIAGTLADKLVHGYLQFYGOEGLSKSVOLNOCPLVMVDGTLPEL 558
DB 432 LFTALNLSNVTQSLQPLQKWSKVDQYLLPHKGDSTLSREYVLDNGLRLQWVDETLPL 491
QY 559 KPRPLPAGRTLVLPVTMGVYVKNVNAALAC 589
DB 492 HEMALAPGSTLGLPAPSGYGVIRNAKAIAC 522

```

```

08T108 PRELIMINARY; PRT; 515 AA.
AC 08T108;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Heparanase-like protein.
GN BHEBPA.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=p50; TISSUE=posterior silk gland;
RA Koike Y., Shimada T., Suzuki M.G., Mita K., Abe H., Maeda S.,
RA Oseogawa K., deJong P.O.;
RT "Genomic sequence of 320kb containing a kettin orthologue on the Z
RT chromosome in Bombyx mori."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB079860; BAB85191.1; "-
DR EMBL: AB090307; BAC10612.1; "-
DR InterPro: IPR005199; Glyco_hydro_79N.
DR Pfam: PF03662; Glyco_hydro_79n; I.
SQ SEQUENCE 515 AA; 59769 MW; FB810ABE6EDDAB CRC64;

Query Match 20.2%; Score 625; DB 5; Length 515;
Best Local Similarity 29.8%; Pred. No. 7.5e-41;
Matches 166; Conservative 90; Mismatches 211; Indels 90; Gaps 17;

QY 60 LILLDVSTKNPVR---TVNENFSLQDPLSIHDGMLDF-----LSSRLV 102
DB 21 MLVLFVKNGGVRFFVINEN---QEDIKLISDFSGIDTIEINRYNSDTRLR 76
QY 103 TLARGLSPAFRFGKRTDFLOFONLRNPAKRGSGPDIYLKNEYEDIVRSVALDKOK 162
DB 77 ELAALSPARLRIGTYSERLIFSKENIPISCH----- 109
QY 163 GCKIAQHPDVLVLOREKAQMLVLLKEQFSNTYSNLILTARSIDLKYNPADCGSLHLI 222
DB 110 NCSYKSYKSLCOLIEKPKCKHKEFL-----PFIIMGEMNQINDFCRTNLTLL 160
QY 223 FALNLRPNPNSSSSALLSLKXSASAKKNISWEI LGNEPNRYRTMGRAVNGSOLGKD 282
DB 161 FSLNMLNRD-NHGMWKKARLEIFSKRKQYALDMQGNENPNSQHPFNBSVTQQLAKD 219
QY 283 YIOLKSLQPIRIYSRASLYGPNIGRP---RKNVIALLDGFMKVASGTDAVWQHCYID 339
DB 220 FEKLRKLLNH-NGYRSHLIVGPDTPRQHPRECLKTIIEFLGNGSHYINRSMHWQYLLN 278
QY 340 GRVYVMDLTKRLLDLSDQIRKIQKQVNTYTPGKTI--WLEBVVTTSAGCTNNLSDSY 397
DB 279 STYALGLEDPFNPEFTDLRQOIETWQNTKTY---KAIPLWLSSTSSSYGGAAGLSNTY 335
QY 398 AAGFLWNTLQMLNOCGIDVIVIRASFPHGYNHLVDQNFNPLDPYWLISLYKRLIGPVL 457
DB 336 AGSPFMIDKLGSAKYNISTVIRQSFIG-GYSLVDEKTLKPLPMMWISLYKRLVGNKVL 394
QY 458 AV--HVAQLQKRPGRVIRDKRLIAAGCTNNHNNHYVAGSTTFTINHSRKXIKLAG 515
DB 395 QVQNCNSRFQ-----RLYTHCNKRYKTNLT-SAVTLVGNLEWAKARFFPLNG 440
QY 516 TL---RDLVHGYLQFYGOEGLSKSVOLNOCPLVMVDGTLPELXPRPLPAGRTLVIP 572

```

Db 441 TALHGDLLIHEYLIA-DSNNKSKTILLNGMPLY--ESNLHINRPIHRYGRVSLP 497  
 Qy 573 PTMGEYVYVNVNALAC 589  
 Db 498 PYSIGFVWIKKISITVC 514

## RESULT 13

Q8H615 PRELIMINARY; PRT; 544 AA.  
 AC Q8H615;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE OSJNBA0035103.11 protein.  
 GN OSJNBA0035103.11  
 OS Oryza sativa (Japanese cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI\_taxid=3947;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 6, BAC  
 clone:OSJNBA0035103."  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003019; BAC2238.1;  
 DR Interpro; IPR005199; Glyco\_hydro\_79N.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 SQ SEQUENCE 544 AA; 58723 MW; ECDE695F0E22A269 CRC64;

Query Match 12.4%; Score 384; DB 10; Length 544;  
 Best Local Similarity 24.8%; Pred. No. 1.3e-21;  
 Matches 163; Conservative 81; Mismatches 198; Indels 216; Gaps 26;

Qy 25 GALYALALLHLSSQAGDRPPLVDRRAAGLKEKTLILDVSTKPFRTVENENFLSLQD 84  
 Db 9 GLCFALL-----RLGGAELAVGVAGVDRRAIATDEDFVCAAMD 50  
 Qy 85 ---PSIHDMDL-----ISSRLVTLARGLSPAL 113  
 Db 51 WMPDCKDYTCGWSGLASLNLVENSFLQLSKFSPPPPPDLSNKLINAIKAFPLKL 110  
 Qy 114 RFGSKRTDFQFQNLKRPAPKSGAP--GPDYLLKVEDDIYRSDVALDKQKCKIAQHPDV 172  
 Db 111 RLGGSLQDKLVY-----GTGDDGGPCAP--FVKN----- 137  
 Qy 173 MLVLQREKAAQMHVLVLEKQFSNTYSNLIITARSIDLKYNFADCGSLHLIFALNALRN- 231  
 Db 138 -----TSEWFGFTQGLPL--HRMDLNAFFQKSGARIVFGINALNGRV 179  
 Qy 232 --FNNS---WSSSALSILKYASKKYNI--SWEIENEPNNYRTMGRAVNSQLGKDYI 284  
 Db 180 PLDDGSGGWDYTNAASTRYASKGYKIHGHEINELSGSGV--GTYAGADQYADVI 237  
 Qy 285 QLSLQPIRIYRSASLYGPNIGRPRKNVIALDGNKVAAGSTVDAVTWQHCYIDGRVYK 344  
 Db 238 ALKSLVD--TIYO-----GNPSKPLVLAAGFEED--AGMFTFVI----- 272  
 Qy 345 VMDFLKTR--LIDPLSDQIRKIQKVNVTYPPGKKI--MLEGVYT-----SAG-- 388  
 Db 273 ---VKTRFNLVAVTHITNLDGVDTHILEKILNPSYIDGNVSTFNSLQGLISAGTS 328  
 Qy 389 -----GTNNLSDSYAAGFLMNTLGLANOGIDIVIRHSFDFHGNHVLVDQ 434  
 Db 329 AVAMVGSAGAYNSGRILVTDSEVFSFWFLDQLGMAKYDTKSYCGQSILGAYGLNKE 388  
 Qy 435 NENPLDPYMLSLYKRLIGPKVLAVHAGIQKRPGRVIRDLRLYACCTHNNHNNYR 494  
 Db 389 TDFQNPDIYLSALMLHRLNGTKVLSATPNG-----TNMRTYAHQAND-----S 431

Qy 495 GSITLPIINL-----HRSKKIKLAACTLDKLVHGYLQPIYQ 532  
 Db 432 PGITILLINLSGNTSGVSYTSEGANNTYKKSRRKTRHIAAGSMRB---EYHLT--AK 485  
 Qy 533 EG-LKSKVQNGOPLVWVDGTLPELKPRPLRAGRTLVIPPTMGFYVYKVNALAC 589  
 Db 486 DGSLSQVMLNGRALVADENGELPRLEPKVDAAQPIAVAPISYVFAHINHPAPAC 543

## RESULT 14

Q9SDA1 PRELIMINARY; PRT; 521 AA.  
 AC Q9SDA1;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN F13G24.30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_taxid=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Beyan M., Van der Schueren J., Chuang Y.J., Voet M., Robben J.,  
 RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Meyer K.F.X.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL133421; CAB62595.1;  
 DR PIR; T45608; T45608.  
 DR GO; GO:0004285; P:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR Interpro; IPR005199; Glyco\_hydro\_79N.  
 DR Interpro; IPR001254; Peptidase\_S1.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hypothetical protein  
 SQ SEQUENCE 521 AA; 57831 MW; 07D8664A4B305CC2 CRC64;

Query Match 12.4%; Score 382; DB 10; Length 521;  
 Best Local Similarity 25.8%; Pred. No. 1.5e-21;  
 Matches 154; Conservative 76; Mismatches 187; Indels 180; Gaps 28;

Qy 75 NENFLSLQD--PSIHD-----GWLDFIS--SKRLVTLA--RGLSPAFIRFGK 119  
 Db 22 DENFVCATLDWMP--HDKNYDCCPWGYSVIMDLTRPLTLKALAFKPLRIRIGSL 78  
 Qy 120 TDFLQFQ--NLNPA-----KSRGGPDPYLLKVEDDIVRSDVALDKQKCKIAQHPDV 173  
 Db 79 QDQYIVDGNLKTCPRFQKNSG-----LFGSKGC----- 110  
 Qy 174 LVLQREKAAQMHVLVLEKQFSNTYSNLIITARSIDLKYNFADCGSLHLIFALNALR- 229  
 Db 111 -----LHM-----KRWDELNSPLTAGAVVITGLNALRREHK 142  
 Qy 230 --FNPNSSWSSSALSILKYASKKYNI--SWEIENEPNNYRTMGRAVNSQLGKDYI 285  
 Db 143 LRGAAGAMOHINTQFQFNVTVYKGVVIDSMWEGNELSGSGV--GASVGAELYKGLDI 200  
 Qy 286 LKSLQPIRIYRSASLYGPNIGRPRKNVIALDGF-----MKAG--STVDATWQH 335  
 Db 201 LKDVIN--KYKSNMLKPLVAP-----GGEYEQQYTKLETSIGSVVDVAVT--HH 249  
 Qy 336 CYIDGR-----VYKWDPLKTRLLDPLSDQIRKIQKVNVTYPPGKKI--MLEGVYT--SAG- 388  
 Db 250 IYNGSGNDPALVYKIND---PSYLSQVSKTFKQVNTQIEHGQ---WASPMVGSAG 302  
 Qy 389 ---GTNNLSDSYAAGFLMNTLGLANOGIDIVIRHSFDFHGNHVLVDQFNPDPYWS 445

Db 303 YNSGRHVSDFIDFWLQDQAGSABHNTKVCQRTLVGGFYGLLEKGTFFVNDYSA 362  
 QY 446 LLYRRLIGPKYLAHVAGLQKRRPGVIRDKLRIYAHCTNNHNNVRSITLFINLH 505  
 Db 363 LLMHRLMGKGLAVQTDG-----PQQLRYAHCSKG-----RAGVTLLINLS 405  
 QY 506 -----RSRKKIKLAGTLRDKLV-----HQYLLQPYGGE 533  
 Db 406 NOSDFTVSVNGINVLNABSRKKSLDITLKRPFWSIGSKASDGYLNREYHLP--EN 463  
 QY 534 G-LKSKSVQNLGQPLVMVDGTLPELKRPLRAGRTLVIPYTMGFYVKNVNALAC 589  
 Db 464 GVLRSKTMVLNKGSLKPTATGDIPLSEPLRBSVNSPILNVLPLMSFTVLPPFDASAC 520

RESULT 15

Q9FF10 PRELIMINARY; PRT: 543 AA.  
 AC Q9FF10;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Similarity to heparanase.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RX MEDLINE=97471969; PubMed=9330910;  
 RA Sato S., Kotani H., Nakamura Y., Kaneo T., Asamizu E., Fukami M.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence  
 RT features of the 1.6 Mb regions covered by twenty physically assigned  
 RT clones."  
 RL DNA Res. 4:215-230(1997).  
 DR ZMBL; AB005249; BAB09947.1;  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR005199; Glyco\_hydro\_79N.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR Pfam; PF03662; Glyco\_hydro\_79n; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 SO SEQUENCE 543 AA; 60250 MM; 0FA2248948282FFE CRC64;

Query Match 12.4%; Score 382; DB 10; Length 543;  
 Best Local Similarity 25.8%; Pred. No. 1.6e-21;  
 Matches 154; Conservative 76; Mismatches 187; Indels 180; Gaps 28;

QY 75 NENPLSLQD--PSLIHD-----GMLDFLS--SKRLVTLA-RGLSPALRFGGR 119  
 Db 44 DENFVCAITLDMWP--HDKCNVQCPWGSVSVINNDLTRPLTKAIRKPLRIRIGSL 100  
 QY 120 TDFLQFC--NLRNPA--KSRGGPDPYLLKNYEDIVRSVDALDKQGGCKIAHPDM 173  
 Db 101 QDQVIYDVGNLKTCPORPFGKNSG-----LFGSKGC----- 132  
 QY 174 LVLDREKAAQMHVLILKEQFSNTYSNLILTRASLDKLYNPADCGSLHIFALNLR----- 229  
 Db 133 -----LHM-----KRWDELNSFLTATGAVVTFGLNALRGRHK 164  
 QY 230 ---RNPNSNNSSSALSLIKTSASKRYNI-SWEIENEPNNRYTMGRAVNGSLQKDYIQ 285  
 Db 165 LRGRAMGAMHINTQDPLANTVSKGYIDSMFENGELSSGV--GASVSAALYKDLIV 222  
 QY 286 LKSLIQPRTIRYSRASLYGNIGRPKNVIALLDG-----MKVAG-STDAVTWQH 335  
 Db 223 LKDVIN--KVYKNSLHKKPIILVAP-----GGFYEQQWYTKLIRISGSPVDVVT--HH 271  
 QY 336 CYIDR-----VYVMDLFTRLDLTSDQIRKIQKVVNTYTPGKKIWLGVVTTTSAG- 388

Db 272 IYNLGSNDPALVKKIMD---PSYLSQVSKTFKVDVNTQIOEHRP-----WASFWGESGGA 324  
 QY 389 ---GTNNISDSYAAGFIMNTLGMLANQIDVYIRHSFEDHGYNHLVDQNFNLPDYLS 445  
 Db 325 YNSGRHVSDFIDFWLQDQAGSABHNTKVCQRTLVGGFYGLLEKGTFFVNDYSA 384  
 QY 446 LLYRRLIGPKYLAHVAGLQKRRPGVIRDKLRIYAHCTNNHNNVRSITLFINLH 505  
 Db 385 LLMHRLMGKGLAVQTDG-----PQQLRYAHCSKG-----RAGVTLLINLS 427  
 QY 506 -----RSRKKIKLAGTLRDKLV-----HQYLLQPYGGE 533  
 Db 428 NOSDFTVSVNGINVLNABSRKKSLDITLKRPFWSIGSKASDGYLNREYHLP--EN 485  
 QY 534 G-LKSKSVQNLGQPLVMVDGTLPELKRPLRAGRTLVIPYTMGFYVKNVNALAC 589  
 Db 486 GVLRSKTMVLNKGSLKPTATGDIPLSEPLRBSVNSPILNVLPLMSFTVLPPFDASAC 542

Search completed: May 6, 2004, 13:45:40  
 Job time: 47 secs



```

Db      708 GGAAGCTCTGAGCAGACACCCCTGACCGACAGCCTGACGCTGACGCTGACGCTG 767
Qy      920 GGGGGCCGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979
Db      768 GGGAGCCCCCGAAGCAGACCCAGCAGCCTGCTCAGAAAGCTTCAAGAAATCTGAGGAG 827
Qy      980 CAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1039
Db      828 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 887
Qy      1040 ACTTCTGAAAACCTGCTGTTGACACACTCTCTGACCAAGATTGAGAAATTCAGAAAG 1099
Db      888 ATTTCCTGAGCCCTGAAAGCTGCTGACCTCTTGGCACTGCACTGCACTGCACTGCA 947
Qy      1100 TGGTAAATACATACATCCCAAGAAAGAAATTTGGCTTGAAGGTGATGATGATGATGAT 1159
Db      948 TCGTGAAGCAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007
Qy      1160 CTGAGGACAGAAACATCTATCCGATTCCTATGCTGACGAGATTCCTATGATGATGAT 1219
Db      1008 GCGGGGGGGGGCCCCAGCTCTCAACACCTATGATGATGATGATGATGATGATGATGAT 1067
Qy      1220 TAGGATGCTGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
Db      1068 TGGGGTTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCT 1127
Qy      1280 GATACATCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1339
Db      1128 GAGCTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1187
Qy      1340 TCTTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399
Db      1188 TATACAAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1247
Qy      1400 AGCCAGGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
Db      1248 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1280
Qy      1460 ACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1519
Db      1281 CCGGGGACCCCAATACCGGGAAGGAGATGATGATGATGATGATGATGATGATGATGAT 1340
Qy      1520 CAGAAAGAAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
Db      1341 TGACCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1400
Qy      1580 TGAAGCTTATGAGGAGAGAGGCTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1639
Db      1401 TGTGCCCCCAGGCAAGGACGATCTGTCAGAGAGGATGATGATGATGATGATGAT 1460
Qy      1640 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1699
Db      1461 TGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1520
Qy      1700 GACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1759
Db      1521 GCAAGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1580
Qy      1760 TGGCTGCTG 1767
Db      1581 TTGCTTGC 1588

```

```

RESULT 2
US-09-435-739-43
; Sequence 43, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Peckel, Itis
; APPLICANT: Viodevsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS

```

```

; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435,739
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 43
; LENGTH: 2396
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-435-739-43

```

```

Query Match      15.9%; Score 282.6; DB 4; Length 2396;
Best Local Similarity 52.8%; Pred. No. 3e-77;
Matches 677; Conservative 0; Mismatches 574; Indels 30; Gaps 2;

```

```

Qy      474 AGATTAACAGAAAGCTGCAAGATTGCGCAGACACCTGATGTTATGCTGCTCCAAAG 533
Db      932 AGTCAACCATGATATTGACAGTCTGAGCCGCTCTGCTGCGTGTGAAGAACTCCA 991
Qy      534 GAGAGAGCAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593
Db      992 GGTGAATAGGCTCTCCAGAGAGCTGTTGCTGCTCCGAGAGCAGTACCAAAAGAGTTCA 1051
Qy      594 TATCTCATATTAACAGCAGCTCTCTAAGCAAACTTTATTACTTTGCTGATGCTG 653
Db      1052 GAAAGCAGCTTACTCAAGAAAGCTCAGTGAATGCTCTACAGTTTGGCAAGTCTGG 1111
Qy      654 ACTCACTGATATTTGCTCTAAATGATGATGATGATGATGATGATGATGATGATGAT 713
Db      1112 GTTAACTGATCTTTGCTCTAAATGATGATGATGATGATGATGATGATGATGATGAT 1171
Qy      714 TTCTAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
Db      1172 CTCAAGCCAGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
Qy      774 ACTGGTATGAGCAATACTATGATGATGATGATGATGATGATGATGATGATGATGAT 833
Db      1232 ACTGGCATAGGCCCAACAGTTTCTGAGAAAGCTCAATCTCTCATGATGATGATGAT 1291
Qy      834 GTTGGAAAGGATTACATCAAGCTGAAAGCTGTTGCAAGCCATCCGATTTATTCAG 893
Db      1292 GTTAGGAAAGACTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1351
Qy      894 AGCCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953
Db      1349 TGCAAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1408
Qy      954 TGAATTCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013
Db      1409 GAGTTTCTGAGGCTGAGGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 1468
Qy      1014 TGAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073
Db      1469 GAATGAGAGATGCTCTCAAAAGAAATTTCTGAGCTGATGATGATGATGATGATGAT 1528
Qy      1074 TGACCAATTAAGAAATTCAGAAAGGTTAATACATCACTCCAGAAAGAAAGATTG 1133
Db      1529 TCTCTGTCGCAAAAATTTGAAAGTCACTAAAGATCAACCTGGAAGAAAGTCTG 1588
Qy      1134 GCTTGAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
Db      1589 GTTGGAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1648
Qy      1194 TGACAGATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
Db      1649 AGCTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1708
Qy      1254 GATACGCACTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
Db      1709 GATAGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1768
Qy      1314 ATTACCAAGTACGCTCTCTCTCTCAAGGCTGATGATGATGATGATGATGATGATGAT 1373

```

Db	1769	TTTACCTGATTA	CTGGGCTCTCTCTTCTG	TCAAGAACTGGTAAG	TCCAGGCTTACT	1828
QY	1374	TGTGCATGTG	CTGGCTGGCTTCCAGCGG	AGGCCAGGCTTGGAGTAT	CCGGCAAACT	1433
Db	1829	GTCAAGATG	TAAGGCCAGA-----	-----	CAGAGCAACT	1861
QY	1434	AAGATTATG	CTCACTGACACAAAC	CACACACCAACTACG	TGTGGGTCCATTA	1493
Db	1862	CCGAGTGTAT	CTCCACTGCACTAA	CGTCTATACCCAGAT	TAACGAAAGATCTAAC	1921
QY	1494	ACTTTTAT	CAATCAACTTGAT	CGATCAAGAAAGAAAT	CAAGCTGCTGGACTCTCAG	1553
Db	1922	TCTGTATGT	CTCGAACTCCATTA	ATGTACCAAGACACTGA	AGTAACGCGCTCGGTGT	1981
QY	1554	AGACAAGT	GTGTTTCAACAGTA	CTGCTGTCACAGGCCAT	TAGGAGAGGGGCTTAAAGTCAA	1613
Db	1982	CAGAAAC	CAAGTGGATGCTAC	TCTTCTAAAGCTT	CGGAGCGGATGATTA	2041
QY	1614	GTCAGTG	CAATGATGGCCAG	CCCTTATGATGTGTG	ACACGGAACCTTCCAGAAAT	1673
Db	2042	ATCTGTCCA	ATGAAACGTCAAA	TTCTGAAGTGTGTG	ATGACGACCTGTCCAGCTTT	2101
QY	1674	GAAAGCCCC	CCCCCTTGGGCGCCG	GCAGATTTGCTATCC	CTCCAGTCAACCATGGGCTT	1733
Db	2102	GACAGAAAA	ACCTCTCCCCCGC	AGAAATGACATAAG	CTGCTCTTTTCTATGATTT	2161
QY	1734	TTATGTG	TCAAGATGTCAA	1754		
Db	2162	TTTTGTCA	TAAAGAAATGCCAA	2182		

```

RESULT 3
US-09-435-739-45
: Sequence 45, Application US/09435739
: Patent No. 6664105
:
: GENERAL INFORMATION:
: APPLICANT: Pecker, Irit
: APPLICANT: Vidavsky, Israel
: APPLICANT: Feinstein, Elena
: TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
: TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
: FILE REFERENCE: 00/20454
: CURRENT APPLICATION NUMBER: US/09/435,739
: CURRENT FILING DATE: 2001-06-05
: NUMBER OF SEQ ID NOS: 47
:
: SOFTWARE: Patentin version 3.0
:
: SEQ ID NO 45
:
: LENGTH: 2396
:
: TYPE: DNA
:
: ORGANISM: Mus musculus
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (594)..(2198)
:
: US-09-435-739-45

```

Query Match	15.9%;	Score 282.6;	DB 4;	Length 2396;
Best Local Similarity	52.8%;	Pred. No. 3e-77;		
Matches 677;	Conservative	0;	Mismatches 574;	Indels 30;
				Gaps 2

QY	474	AGATTAACAGAAAGGCTGCAGATTTGCCAGACACCTGATGTTATGTCGTCTCCAAAG	533
Db	932	AGTCACCACTGATTTTGCAGGCTCGAGACGGGCTCTGCTGGGGGTGTAGGAAACTCCA	991
QY	534	GGAGAAAGCAGCTCAGATGCACTGTGGTTCTTTAAAGAGCAATTCCTCAATCTTACAG	593
Db	992	GGTGAATAGCCCTTCCAGAGCTGTGTGCTCCGAGGCAAGTACCAAAAGAGTCCA	1053
QY	594	TAATCTCATATTAAACAGCCAGGCTCTCAAGCAAACTTTTAATCTTGGTGAATTCGTCTGG	653
Db	1052	GAAACAGACCTTACTAGAGAGCTCAGGTGACATGCTCTTCACTTTTGGCCAAAGTGGTCGGG	1111
QY	654	ACTCACTGATTTTGTCTCTTAAATGCACTGGGTGTAAATCCCAATAACTCTCTGGACAG	713

Db	1112	GTTCAGACCTATCTTTGGTCTTAATGGCTTACTACGAACCCCAAGACTTACGGGTGAACAG	1177
Qy	714	TTTCATAGCCCTGAGCTGTGTGAAGTACAGCGCCAGCAAAAAGTACAACATTTCTTGGGA	773
Db	1172	CTCCACGCCCCAGCTTTCTCTTGACTGAATGCTCTTCCAAAGGTTATACATCTTCCTGGGA	1231
Qy	774	ACTGGGTATATAGCCCAATATCTATCCGACCATGCAATGCGCGGACGTAAATGGCAGCCA	833
Db	1232	ACTGGGCAATGAGCCCAACAGTTTCTGGAGAAAGTCACTTCTCATTCGATGGGTGCA	1291
Qy	834	GTTGGGAAAGGATTACATCCAGCTGAAGAGCTGTGGACCCATCCGATTTATTCAG	893
Db	1292	GTTAGGAAAGACTTGTGGAGTTGCAATTAATTCTTCA--AAGGTACGTTTCCAAA	1348
Qy	894	AGCCAGCTTATATGGCCCTAATATTTGGCCGCGCAGGAGAAATGTCAATGCCCTCTTAGA	953
Db	1349	TGCAAAACTCTATGCTCTGACATCGGTCAAGCCTCGAGGGAAAGACATTTAACTGCTGAG	1408
Qy	954	TGGAATTCATGAAGGTGGCAGGAAGTACAGTATGATGCACTTAACTCTGGCAACATTGCTCAT	1013
Db	1409	GAGTTTCTTGAAAGGCTGCGCGGAGAGTATCGATCTCTTACATAGGGATCATATTACTT	1468
Qy	1014	TGATGGCCCGGGTGTCAAGGTATGACCTTCCTGAAAACTCGCCTGTTAGACACACTCTC	1073
Db	1469	GAAATGAGATCATGCTTACCAAGAAAGATTTCTGAGCTCTGATGCGCTGGACACTTTTAT	1528
Qy	1074	TGACCAGATTAGAAAAATTCAGAAAGTGGTTAATACATACATTCACAGAAAGAAATTTG	1133
Db	1529	TCTCTGTGTGCAAAAAATTCGAAAGTCACTTAAGAGATCACCTGGCAAGAAAGTCTG	1588
Qy	1134	GCTTGAAGAGTGTGGTACACACTCAGCTGGAGGCAAAACAATATATCCATTCCTATGC	1193
Db	1589	GTTGGGAAGAAGAGCTCAAGCTTACGGGTGGGTGACCTCTGCTGTGCACACCTTTGTC	1648
Qy	1194	TGCAGATTTCTTATGTTGTGAACACTTTTAGAATCTGTGCCAATCAGGCAATTGATGTCGT	1253
Db	1649	AGCTGGCTTTATGTGCTGGGTAATATTTGGCCTGTTCAGCCCAAGATGGCAATAGAATGCT	1708
Qy	1254	GATACGCGCATCATTTTTTTGACCATGGATACATATCACTCGTGGACCAAGATTTTAACC	1313
Db	1709	GATGAGCGAGGTGTTCCTTCGAGAGAGGACATACACTATATGATGTAAGAACTTTGAGCC	1768
Qy	1314	ATTACCAAGACTAGGGCTCTCTCTCTCTCAAGAGGCTGATCGAGCCCAAGTCTTGGC	1373
Db	1769	TTTACTCTATTAATCGCTCTCTCTCTCTTTCTGTTCAAGAAACTGGTAACTCCACAGGTGTTACT	1828
Qy	1374	TGTGATSTGGCTGGGCTCCAGCGGAAGCCAGCGCTGGCCGAGTATCCGGACAAACT	1433
Db	1829	GTCAGAATGTAAAGGCCAG-----CAGAGCAAACT	1861
Qy	1434	AAGATTTATGCTACATCGACAAACCAACAACAACAACACTAGTGGTGGTGCATATAC	1493
Db	1862	CCGAGTATATCTCACTGACCTAACGTTATACACCACAGATATACAGGAAGAGATCTAAC	1921
Qy	1494	ACTTTTATCATCAACTTGATCGATCAAGAAAGAAATCAAGCTGGCTGGGACTCTCAG	1553
Db	1922	TCTGTATGTCTGAACCTCCATATATGTACACAGACATTTGAAGTACCGCTCCGTGTGT	1981
Qy	1554	AGACAACTGTTTACACAGTACTCTGCTGCACCTCTATGGGAGAGAGGCCCTAAAGTCAA	1613
Db	1982	CAGGAAACAAATGATAGTACGTACTTGTGAACCTTCGGGAGCGAGTGAATTAATTTCCAA	2041
Qy	1614	GTCAGTGAATGGAATGGCCAGCCCTTAGTATGATGGTGGACGACGGGACCTCTCCAGATTT	1673
Db	2042	ATCTGTCCAATGAAAGGTCTAAATTTGAAATGTTGATAGAGACACCTGCGCAGCTTT	2101
Qy	1674	GAAAGCCCGCCCTTGGGCGCGGCGGACATTTGTCATTCCTTCAGTCAACATGGGCTT	1733
Db	2102	GACAGAAAAACCTTCCCCCGAGAGAGTCACTAAGCCTGCTGCTTTTCTATGTGTTT	2161
Qy	1734	TTATGTGTCAGAATGTCAA	1754
Db	2162	TTTTGTCAATGAAGATGCCAA	2182



RESULT 4  
US-08-922-170B-9  
Sequence 9, Application US/08922170B  
Patent No. 5968822  
GENERAL INFORMATION:  
APPLICANT: Iris Pecker, Israel Vlodavsky and Elena  
APPLICANT: Feinstein  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
TITLE OF INVENTION: HAVING HEPARANASE ACTIVITY AND EXPRESSION OF  
TITLE OF INVENTION: SAME IN TRANSFUSED CELLS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein  
STREET: 2940 Birchtree lane  
CITY: Silver Spring  
STATE: Maryland  
COUNTRY: United States of America  
ZIP: 20906  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead Slimote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
OPERATING SYSTEM: Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to  
SOFTWARE: an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/922,170B  
FILING DATE: 2 SEP 1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-5625553  
TELEFAX: 972-3-5625554  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-922-170B-9  
Query Match 15.7%; Score 278.8; DB 2; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;  
QY 474 AGATTAACAAAGAGGCTGCAAGATGCGCCAGCAGCCCTGATGTTAGCTGCTGCTCAAG 533  
DB 425 AGTCAACAGAGATTTTGGCAATATGATGATCCCTCTGATGAGAGAGAGATTAG 484  
QY 534 GGAAGAGGAGCTGATGATGATGCTGTTCTTAAAGAGCAATTTCCAACTACTTACAG 593  
DB 485 GTTGAATGAGCCCTCAACAGAGATTTGCTACTCCGAGAACTACCAAGAAAAGTTCA 544  
QY 594 TAATTCATTTAAACAGCCAGGCTCTAGCAAACTTTTAACTTTGCTGATTTGCTGAG 653  
DB 545 GAACAGACCTACTCAAGAGAGCTCTGATGATGCTATACACTTTTGCAAACTGCTCAG 604  
QY 654 ACTCACTGATATTTGCTTAAATGACATGCTGATATCCCAATTAACCTCTGAGAACAG 713  
DB 605 ACTGAGCTTATCTTTGCTTAAATGATGCTTATTAAGACAGCAATTTGAGGAGAACAG 664  
QY 714 TTCTAGGCTTGAAGTGTGTTGAAGTACAGGCGCAAGAAAAGTACAACTTTCTTGGGA 773  
DB 665 TTCTAATGCTCAAGTGTGCTCTGAGACTAGCTCTTCCAAAGGGGTATTAACATTTCTTGGGA 724

RESULT 5  
US-08-922-170B-11  
Sequence 11, Application US/08922170B  
QY 774 ACTGGTATATGAGCAAAATTAATATGAGCAATGATGAGCGGAGAGTAAATGACAGCA 833  
DB 725 ACTAGGAGATGAGACCTTAACAGTTTCTTAAAGAGGCTGATATTTTATCAATGAGTGGAG 784  
QY 834 GTTGGGAAAGATTAATCACTGAGCTGAAAGCTGTTGACGCCATCCGATTTATTCAG 893  
DB 785 GTTGGGAAAGATTAATTAATGATTAAC---TTCTAAGAAAGTCACTTCAAAA 841  
QY 894 AGCCAGCTTATGAGCCCTTAATATGAGCGGCGGAGAGAAATGATGATGCGCCCTCTAG 953  
DB 842 TGCAAACTCTATGCTCTATGTTGCTGAGCTCGAAGAAACGCTTAAGTGTGA 901  
QY 954 TGATTCATGAGGCTGAGAGAGTACAGTATGATGATGATGATGATGATGATGATGAT 1013  
DB 902 GAGCTTCTGAGAGCTGATGAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 961  
QY 1014 TGATGCGCGGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
DB 962 GAATGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
QY 1074 TGACCAATTAAGAAATTAAGAAAGTATTAATTAATTAATTAATTAATTAATTAAT 1133  
DB 1022 TTCACTGTGCAAAAAGTTTCCAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081  
QY 1134 GCTTGAAGGCTGATGAG 1193  
DB 1082 GTTGGGAAAGCAAGCTGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141  
QY 1194 TGCAAGATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
DB 1142 AGCTGCTTATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201  
QY 1254 GATAGGAGCTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313  
DB 1202 GATAGGAGATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1261  
QY 1314 ATTAACAGACTACTGAGCT 1373  
DB 1262 TTACCTGATTTATGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321  
QY 1374 TGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1433  
DB 1322 GCGAAGCTGCAAGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1354  
QY 1434 AAGATTTATGCTCACTGAGCAAAACCAACCAACCAACCAACCAACCAACCAACCAAC 1493  
DB 1355 TCGATATCTCTATTTGCAACAACTGATGATGATGATGATGATGATGATGATGATGAT 1414  
QY 1494 ACTTTTATATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553  
DB 1415 TCTGATGCTCAATTAACCT 1474  
QY 1554 AGACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613  
DB 1475 TAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1534  
QY 1614 GTCAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673  
DB 1535 ATCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594  
QY 1674 GAAGCCCGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1733  
DB 1595 AATGAGAAACCTCTCTGCGGCGGAGAGATGATGATGATGATGATGATGATGATGAT 1654  
QY 1734 TTATGCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
DB 1655 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1700

Patent No. 5968822  
 GENERAL INFORMATION:  
 APPLICANT: Iris Pecker, Israel Vlodavsky and Elena  
 APPLICANT: Feinstein  
 TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
 TITLE OF INVENTION: HAVING HEPANASE ACTIVITY AND EXPRESSION OF  
 TITLE OF INVENTION: SAME IN TRANSFUSED CELLS  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein  
 STREET: 2940 Birchtree lane  
 CITY: Silver Spring  
 STATE: Maryland  
 COUNTRY: United States of America  
 ZIP: 20906  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: Twinhead\* Siliemote-890TX  
 OPERATING SYSTEM: MS DOS version 6.2,  
 OPERATING SYSTEM: Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to  
 SOFTWARE: an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/922,170B  
 FILING DATE: 2 SEP 1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-922-170B-11

Query Match 15.7%; Score 278.8; DB 2; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

QY 474 AGATTAACAGAAAGGCTGCAAGATTGGCCAGCAACCTGATGTTATGCTGCTCCAAAG 533  
 DB 425 AGTCAACCAAGATATTTGCAAAATATGATCCATCCCTCGATGTGAGAGAAAGTTAG 484  
 QY 534 GAGAAAGGACGCTCAGATGCACTGCTCTTTCTTAAAGAGCAATTTCTCCATATCTTACG 593  
 DB 485 GTTGGAATGGCCCTTACCAAGAGCAATTTCTTACCTCCAGAACACTACAGAAAGTTCAA 544  
 QY 594 TAATCTCATATTAACAGCCAGCTCTCTAGACAACTTATTAATCTTGTGCTATGCTCTG 653  
 DB 545 GAACAGACCTACTCAAGAAAGCTCTGTAGATGTGCTATTAACCTTTTGCAAACTGCTCAG 604  
 QY 654 ACTCCACTGATATTTGCTCTTAATGCACTGCGCTGTATCCCAATTAATCTCTGAAACG 713  
 DB 605 ACTGCACTTATCTTTGGCTTAATATGCTTATTAAGAACAGCAATTTGCAAGTGAACAG 664  
 QY 714 TTCTAGTGCCTGATGCTGTTGAAGTACAGGCGCAAGAAATTAACAATTTCTTGGGA 773  
 DB 665 TTCTATGCTCAAGTGTCTCTGCACTACTCTTCTTCAAGGGTATTAACAATTTCTTGGGA 724  
 QY 774 ACTGGGTAATGAGCCAAATTAATCTATGCAATGATGATGCGCGGAGTAATGAGCA 833  
 DB 725 ACTAGGCAATGAAGCTTAACAGTTTCTTAAAGAGGCTGATATTTTATCAATGGGTGCA 784

QY 834 GTTGGAAAGATTATACATCCAGCTGAAGAGCTGTGGAGCCCATCCGATTTATTCAG 893  
 DB 785 GTTAGAAGAAATTAATATTCATATGCAATTAAC---TTCTAAGAAAGTCCACTTCAAAA 841  
 QY 894 AGCCAGCTTATATGCTCCCTAATTTATGGCGGCGGAGAGAAATATCATGCCCTCTAGA 953  
 DB 842 TGCAAACTCATATGCTCTGTATGTTGGTCAAGCTTCAAGAAAGAGCGCTAAGATGCTGAA 901  
 QY 954 TGATTCATGAAGGAGGAGAGAAATGACAGTATGACAGTATCTGCAACATTTCTCAT 1013  
 DB 902 GAGCTTCTGAAAGCTGTGTGAGAGAGATGATTAATCACTTACATGCACTACTCTATT 961  
 QY 1014 TGATGCGGCGGTGTGCAAGTGTATGAGACTTCTGTAATACTGCTGTATGACACACTCTC 1073  
 DB 962 GAATGAGCGACTGCTTACAGAGAAAGATTTTCTAACCCTGATGATATGCAATTTTAT 1021  
 QY 1074 TGACCAATTTGAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGAAATTTG 1133  
 DB 1022 TTCACTGTGCAAAAGTTTTCAGGTGTGTGAGAGACACAGGCTTGCAAGAAAGTCTG 1081  
 QY 1134 GCTTGAAGTGTGTGACCACTCAGCTGAGGAGCAAAACATCTATCCGATTTCTATGC 1193  
 DB 1082 GTTAGAAGAAACAGCTGTGATATGAGAGCGGAGCGGCGCTGTATCCGACACTTTGC 1141  
 QY 1194 TGCAAGATTTTATGTTGAAACACTTTAGAAATGCTGGCCATCATGAGGATTTATGCT 1253  
 DB 1142 AGCTGGCTTTATATGCTGTGATTAATTTGGGCTTCAAGCTTCAAGCTTCAAGTGAATG 1201  
 QY 1254 GATACGCACTCATTTTGTGACATGATACATCACTCTGTGACCAAGATTTTAAACC 1313  
 DB 1202 GATGAGCAAGATTTCTTTGAGAGAGAACTACATTTATGATGATGATGATGATGATG 1261  
 QY 1314 ATTACCAACTATGCTGCTCTCTCTCTCTTACAGGCTGTATGCGCCCAAGTCTTGC 1373  
 DB 1262 TTATCTATTTATGCTGCTATCTCTCTCTCTTCAAGAAATTTGCTGAGCACTCAAG 1321  
 QY 1374 TGTGATGTGGCTGGCTCCAGCGAAGCCAGGCTGTGAGTATCCGGAGCAAACT 1433  
 DB 1322 GGCAAGCTGCAAGTTTCAAGAGAGG-----AGCT 1354  
 QY 1434 AAGATTTATGCTCATCTGACAAACCAACACACAACTACCTGTGCTGCTCATTTAC 1493  
 DB 1355 TCGAGTATACCTTCACTGACAAACACTGACAACTCAAGATTAAGAGAGATTTTAC 1414  
 QY 1494 ACTTTTATCACTCACTGATCGATCGATCAAGAAATCAAGCTGTGAGATCTTCAG 1553  
 DB 1415 TCTGATCCCAATTAACCTCCATTAACAGTCAACCAAGTATCTGCGTATCTATCTTTTC 1474  
 QY 1554 AGACAGCTGTTTCACTGACAGTACCTGTCAGCCCTATGAGGAGAGGCTTAAAGTCCA 1613  
 DB 1475 TAACAGAGAGTGAATTAATCTTTCTTAAGACCTTTGGGACCTCATGATTAATCTTCAA 1534  
 QY 1614 GTCACTGCACTGAATGAGCCAGCCCTTATGATGTGTGAGACAGGAGACCTTCCAGAA 1673  
 DB 1535 ATCTGTCACTCAATGATGCTTAAATATGATGATGATGATGATGATGATGATGATGAT 1594  
 QY 1674 GAAGCCCGCCCTTCCGAGCGGCGGAGCAATTTGTATCTCTCAGTCAACATGAGGCTT 1733  
 DB 1595 AATGAAAAACCTCTCCGCGCAGAGAGTTCACTGGGCTTGCACTTTCTCATATAGTTT 1654  
 QY 1734 TTATGTGTCAAGATGTCATGCTTTGGCTGTGCGCTGACCGATTA 1779  
 DB 1655 TTTTGTGATGAAGAAATGCAAGTGTGCTGTGATCTGAAATTA 1700

RESULT 6  
 US-09-071-739B-1  
 Sequence 1, Application US/09071739B  
 Patent No. 6177345  
 GENERAL INFORMATION:  
 APPLICANT: Iris Pecker et al.  
 TITLE OF INVENTION: HEPANASE SPECIFIC MOLECULAR PROBES  
 TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL



CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimnote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
SOFTWARE: Word for Windows version 2.0 converted to  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,7398  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33, 883  
REFERENCE/DOCKET NUMBER: 910/5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: Nucleic acid  
STRANDEDNESS: Double  
TOPOLOGY: linear

US-09-071-7398-3

Query Match 15.7%; Score 278.8; DB 3; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

QY	474	AGATAAACAAGAGCTGCAAGTTCGCCAGACCCCTGATGTTATGCTGCTGCTCAAG	533
DB	425	AGTCAACCGAGATATTTGCAAAATATGATCAATCCCTCTGATGAGGAGAAATTCG	484
QY	534	GAGAGAGGAGCTGCAATGATGCTGTTCTTAAAGAGCAATCTCCATTAATTAAG	593
DB	485	GTTGATGAGCCCTACAGAGCAATTCCTACCGAGAACTACAGAAAAGTTCAA	544
QY	594	TATCTCATATTAACAGCAGCTCTTAGACAACTTTATTAATTGCTGATTCCTGG	653
DB	545	GAAAGACACTACTCAAGAGCTCTGATGCTATACCTTTTGCAAACTGCTCAGG	604
QY	654	ACTCCACCTGATTTGCTTAAATGACCTGCTGTATACCCATTAATCCTCGAAGC	713
DB	605	ACTGACTGATCTTTGGCTTAAATGCTTATTAACAAGCAAGTTGCACTGAGAACG	664
QY	714	TTCTATGCTGCTGAGCTGTTGAGTACAGCCGCAAGCAAAAGTACACATTTCTGGG	773
DB	665	TTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	724
QY	774	ACTGGGTATGAGCCCAATTAATCTATGACCATGCTGCGGCGAGTAAATGGCAGC	833
DB	725	ACTAGGCAATGACCTTAAGATTTCTTAAAGAGCTGATTTTCATCAATGGCTGCA	784
QY	834	GTTGAGAAAGATTAATCACTGAGTGAAGCTGTTGACAGCCATCCGATTTATTCAG	893
DB	785	GTTAGAGAAAGATTAATCACTGATTAAC---TTTAAAGAAATCCACCTTCAAAA	841
QY	894	AGCCAGCTATATGCGCTTAATTTGGGCGCGCGAGAGAGATGTATGCGCTCTTAA	953
DB	842	TGCATAAATCTATGCTGCTGATGTTGTTGCTGAGCTTCAAGAAAGACGCTTAAG	901
QY	954	TGATTCATGAAGGTGAGAGAGTACAGTATGAGTACCTGCGCAATTCGTACAT	1013

DB	902	GAGCTTCCTGAAGCTGAGAGAAAGTATTCAGTTACATGATCCTACTATTT	961
QY	1014	TGATGCGCGGCTGTCAGAGTATGAGACTTCTGAAAATCTCGCTGTTAGACACTCTC	1073
DB	962	GATGAGACGAGCTGTCACAGGAGAAATTTTCTAAACCTGATGATTTGACATTTTAT	1021
QY	1074	TGACCAATTTAGAAAATTCAGAAATGCTTAATATACACTCCAGAGAAAGATTTTG	1133
DB	1022	TTATCTGTCGCAAAAATTTTCCAGGTGTTAGAGCACAGCCCTGCGCAAGAGTCTG	1081
QY	1134	GCTTGAAGGTGTGTACCACTCAGCTGAGGCAACAACTATATCCATTCCTATGC	1193
DB	1082	GTTAGAGAAAACAAGCTCTGATATGAGAGCGAGCCCTTGCTATCCGACCTTGC	1141
QY	1194	TGACAGATTTTATGCTTGAACCTTTAGAAATGCTGCGCAATCAGGCTATGATGCT	1253
DB	1142	AGCTGCTTTATGCTGCTGATTAATTTGGCCCTGTCAGCCCGCAATGGAATGAGTGT	1201
QY	1254	GATAGGCACTATTTTGGTACATGATATCAATCACTCTGAGACCAATTTTAACC	1313
DB	1202	GATGAGCAAGATTTCTTGGAGAGAACTACCAATTAATGATGAAATCTTCATTC	1261
QY	1314	ATTACCACTACTGCT	1373
DB	1262	TTTACCTGATTTATGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1321
QY	1374	TGTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1433
DB	1322	GCGAAGCTGCAAGTTCAGAAAGAGAGG-----AAGCT	1354
QY	1434	AAGATTTATGCTACTGCAACAACCAACCAACCACTACGTTGCTGCTCATTAAC	1493
DB	1355	TGAGATTAATCTTATGCAACAACCACTGCAATCAAGTATTAAGAGAGATTTTAC	1414
QY	1494	ACTTTTATCATCACTTGCATGATCAAGAAAGAAATCACTGCTGCTGCTGCTG	1553
DB	1415	TTCTATGCTATTAACCTCCATTAACGTCACCAAGTCTTGGCTTATCCCTATCTTTC	1474
QY	1554	AGAAAGCTGTTCAACAGTACCTGCTGCAAGCCCTTATGAGGAGAGAGGCTTAAGTCCA	1613
DB	1475	TACAGAGAGAGATTAATTAATCTTAAAGCTTTGGAGCTCTGATGATTAATCTTCCA	1534
QY	1614	GTCAGTCACTGATGAGGAGCCCTTATGATGATGATGATGATGATGATGATGATG	1673
DB	1535	ATGTCCTCACTCATGCTTAACTTAAGATGATGATGATGATGATGATGATGATGAT	1594
QY	1674	GAAAGCCGCGCTTGGGCGCGCGGAGATGATGATGATGATGATGATGATGATGATG	1733
DB	1595	AATGAAAACCTCTCGGCGAGAGAGTCACTGAGGCTTCCAGCTTCTCATATAGTTT	1654
QY	1734	TTATGCTCAAGATGCTATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1779
DB	1655	TTTGTATTAAGAAATCCAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1700

RESULT 8  
US-09-260-0388-1  
Sequence 1, Application US/092600388  
Patent No. 6348344

GENERAL INFORMATION:  
APPLICANT: Nely Ayal-Hershkovitz et al.  
TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
EXPRESSING RECOMBINANT HEPARANASE  
AND METHODS OF PURIFYING SAME

NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
STREET: 2001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: Twinhead\* SImote-890TX  
 OPERATING SYSTEM: MS DOS version 6.2,  
 Windows version 3.11  
 SOFTWARE: Word for Windows version 2.0 converted to  
 an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/260,038B  
 FILING DATE: 02-Mar-1999  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedmam, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-260-038B-1

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

QY 474 AGATTAACGAAAGGCTGCAAGTTCACCAACCTGTATGTTATGCTGCTCCAAAG 533
DB 425 AGTCAACGAGATTTGGCAATATGATCCATCCCTCTGATGAGGAGAAAGTTACG 484
QY 534 GAGAGAGAGCTCAGATCATCTGCTTCTTAAAGAGCAATTCCTCAATCTTAACG 593
DB 485 GTTGAATGGCCCTACAGAGAGAAATGCTACTCCAGAGACACTACAGAAAAATTGCA 544
QY 594 TAATCTCATATTAACAGCCAGGCTCTAGACAACTTTTAACTTTGCTGATGCTGCG 653
DB 545 GAACAGACACTTCTCAAGAGCTGTAGATGCTATACCTTTTGCAACCTGCTCAGG 604
QY 654 ACTCACTGATATTTGCTCTAATATGCTGAGCTGATCCCAATTAATCTCTGGAACG 713
DB 605 ACTGACTGATCTTTGGCTTAATGCTTATTAAGAACAGCAATTTGCACTGGAACG 664
QY 714 TTCTAGTCCCTGAGTCTGTTGAAGTACAGCCGCAAGAAAAAGTACAACTTTCTTGGA 773
DB 665 TTCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724
QY 774 ACTGGTATGAGCCAAATTAATCATGAGCACTGATGAGCCGGGAGTAAATGAGCAGCA 833
DB 725 ACTAGCAATGAACTTAAAGCTTTCTTAAAGAGCTGATATTTTCAATCAATGAGTCCA 784
QY 834 GTTGGAGAAAGATTAATCATCACTGAGTGAAGCCTGTTGAGCCCACTCCGATTTATTCAG 893
DB 785 GTTAGAGAGAAAGTATTAATCAATGATCAAAAC---TTCTAAGAAAGTCCACCTTCAAAA 841
QY 894 AGCAGACTATATGCGCTTAATTTGGGGGGGCGAGAGAAATGTCATGCGCTCCTAGA 953
DB 842 TGCAAACTCTATGCTGCTGATGTTGATGAGCTTCAAGAAAGAGCGCTTAAGTGTGA 901
QY 954 TGATTCATGAAGGTGAGAGAGTACAGTAAATGAGTACCTTACCTGCAACATGCTACAT 1013
DB 902 GAGCTTCTGAAAGGCTGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 961

```

```

QY 1014 TGATGCGCGGTGCTCAAGGTATGAGCTTCTGAAAAGTCCGCTGTAGACACTCTC 1073
DB 962 GAATGAGCGAGCTGCTACCAAGAGAAATTTCTTAAACCTGATGATTTGGAATTTTAT 1021
QY 1074 TGACAGATTAAGAAATTCAGAAAGTGTATATACATCACTCCAGAGAAAGATTGG 1133
DB 1022 TTATCTGTGCAAAAAGTTTCCAGGTGTTGAGACCAAGGCTGCGCAAGAGGTCTG 1081
QY 1134 GCTTGAAGGTGTGTACCACTCAGCTGAGCAACAAACATTCATTCGATTCATGCG 1193
DB 1082 GTTAGAGAGAAACAAGCTTGCATATGAGAGCGAGCGCTTGTATTCGACACTTTGCG 1141
QY 1194 TGACAGATTTCTATGTTGTTGAACCTTTAGAAATGCTGCGCAATCAAGGCAATGATGCTG 1253
DB 1142 AGCTGGCTTTATGCTGCTGATTAATTTGGGCGCTGAGCCGCAATGGAATGAGATGCT 1201
QY 1254 GATACGCACTCATTTTGGTACCATGATACATCACTCTGTGACCAAGATTTTAAACC 1313
DB 1202 GATGAGGCAAGTATTTCTTTGAGACAGAAACTACATTTAGTGAAGAAACTTGCATCC 1261
QY 1314 ATTACGAGACTACCTGCTCTCTCTCTACAAAGCGCTGATGCGCCCAAGTCTGCGC 1373
DB 1262 TTATCTGATTTATGCTGCTATCTTCTGTTCAAGAAATTTGTTGGGACCAAGGTGTTAT 1321
QY 1374 TGTGATGTGCTGGGCTCAGCGGAGCCAGCGCTGCGCAGATGATCCGGACAAACT 1433
DB 1322 GGCAGAGCTCAAGTGTCAAGAGAGAG-----AAGCT 1354
QY 1434 AAGGATTTATGCTCAGTGCAGCAACCAACCAACCACTAGTGTGCTGCTATAC 1493
DB 1355 TCGAGTATACCTTATGCTGCAAAACACTGACAAATCCAGTATTAAGAGAGATTTAAC 1414
QY 1494 ACTTTTATCATCACTTGCATGATCAAGAAAGAAATCAAGTGGCTGAGACTCTCAG 1553
DB 1415 TCTGTATGCACTAACTCATATACGTACCAAGTCTTGGCTTACCTTATCCTTTTTC 1474
QY 1554 AGCAAGCTGTTCAACAGTACCTGCTGAGAGCCCTTGGGAGAGAGGCTTAAAGTCCA 1613
DB 1475 TAAACAGAGTGTATTAATTAATCTTTTAAGACCTTTGGAGCTCAGATGATCTTCCAA 1534
QY 1614 GTCAGTCAACTAATGAGCCAGCCCTTATGATGTGAGAGAGAGAGCCCTCCAGAAAT 1673
DB 1535 ATCTGCTCACTCAATGATGTCTAATCTTAAAGATGTGATGATCAACCTTGCACCTTT 1594
QY 1674 GAAGCCCGCCCTTCCGCGCGCGGAGCATGTGATCATCCCTCCAGTACCAATGGGCTT 1733
DB 1595 AATGAGAAACCTCTCCGCGCAGAGAAATTCATGAGGCTTGCAGCTTTCTCATATAGTTT 1654
QY 1734 TTATGTGTCAGAAATGTCATGCTTTGGCTGCGCTACCGCTAACGATTA 1779
DB 1655 TTTTGATTAAGAAATGCCAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700

```

RESULT 9  
 US-09-260-038B-3  
 ; Sequence 3, Application us/09260038B  
 ; Patent No. 6348344  
 ; GENERAL INFORMATION:  
 APPLICANT: Maty Ayal-Herszkovitz et al.  
 TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
 EXPRESSING RECOMBINANT HEPARANASE  
 AND METHODS OF PURIFYING SAME  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead\* Slimnote-890TX  
 OPERATING SYSTEM: MS DOS version 6.21  
 SOFTWARE: Word for Windows version 3.11  
 an ASCII file  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/260,038B  
 FILING DATE: 02-Mar-1999  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/071,618  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedmann, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/16  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-260-038B-3

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

QY 474 AGATAAAGAAAGGTCAGAGTTGCCAGACCTCTGATGTTTGTCTGTCTGCTCCAAAG 533
DB 425 AGCAACACAGAGATATTTGCAATATGATTCATCTCTGATGTCAGAGAAAGTTGCG 484
QY 534 GGAGAGGAGAGCTCAGATGATCTGTTCTTAAAGAGCAATTCATTAATCTTACAG 593
DB 485 GTTGAGATGGCCCTACAGAGAGCAATGCTACTCCAGAAACCTACCAAAAAAAGTCA 544
QY 594 TAATCTCATTTTAAAGCCAGAGCTCTAGACAACTTTATPACTTTGCTGATGCTGG 653
DB 545 GAACAGACCTTACTCAAGAGAGCTGTAGATGTCTATACCTTTTGCACAACTGCTCAG 604
QY 654 ACTCCAGCTGATTTTGTCTCTAAATGCACTGGGTGTATTCACCAATPACTCCTGAGAG 713
DB 605 ACTGGACTTGATTTTGGCCCTTAATGCTTTATTAAGAACAGAGATTTGCAAGTGAACG 664
QY 714 TTCTAGTCCCTGAGTCTGTTGAAGTACGCGCAGCAAAAAGTACAACTTTCTTGGGA 773
DB 665 TTCTATATGCTAGTGTCTCTGACTAGCTCTTCCAGAGGATTAATCAATTTCTTGGGA 724
QY 774 ACTGGGTATGAGCCAAATTAAGTACAGCAATGATGAGCCGAGGAGTAAATGAGAGCA 833
DB 725 ACTAGCAATTAAGCTTAACAGTTTCTTAAAGAGGCTGATATTTTCATCATATGAGTGGCA 784
QY 834 GTTGGAGAGATTAATCATCAGCTGAAGAGCTGTTGACGCCCATCCGATTTATTTCCAG 893
DB 785 GTTAGAGAGATTAATTAATTAATGATTAAC---TTTAAAGAGTTCACCTTCAAAA 841
QY 894 AGCCACTTATATGAGCCCTAATATTTGGGGGCGAGAGAGATGTCATCCCTCTTACA 953
DB 842 TGCAAACTCTATAGTCTGATGTTGTGTAAGCTCAGAAAGACGGCTTAAGATGTGA 901
QY 954 TGGATTCATGAAGGTGAGAGAGTACAGTATGATGATTAATGATTAATGATTAAT 1013
DB 902 GAGCTTCGTAAGGCTGTGAGAGAGTATGATTAATGATTAATGATTAATGATTAAT 961

```

```

QY 1014 TGATGCGCGGATGTCAGAGTATGACTTCTGAAAACCTGGCTGTAGACACTCTC 1073
DB 962 GAATGACGAGCTCTACAGAGAAATTTCTTAACCTGATGATTTGACATTTTAT 1021
QY 1074 TGACCAATTAAGAAAATTCAGAAAGTGTAAATACATACCTCCAGAAAAGATTGG 1133
DB 1022 TTCACTGTGCAAAAATTTTCCAGAGGTGTGAGACACACAGGCTCGCAAGAGCTCG 1081
QY 1134 GCTTGAAGGTGTGTGACCACTCAGCTGAGAGCAAAAACATCTATCCGATTTCTATGC 1193
DB 1082 GTTAGAGAAAACAGCTCTGATATGAGAGGAGGAGGCTTGTCTATCCAGCTTTTGC 1141
QY 1194 TGACAGATTTCTTAATGTTGAACCTTTAGAAATGTCGGCCATCAGGCAATGATGCT 1253
DB 1142 AGCTGGCTTATATGAGCTGATTAATTTGGGCTGTGAGCCGATGGAATTAAGATGTG 1201
QY 1254 GATACGGCACTATTTTTCATGATGATATCATCATCTGTCGACCAAGATTTTAACC 1313
DB 1202 GATAGGCAAGATTTCTTGAAGAGAACTACATTTAGTGAAGAACTTCATCCATCC 1261
QY 1314 ATTACGACTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1373
DB 1262 TTACCTGATTTATGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
QY 1374 TGTGATGTGCTGGCTCTCCAGCGGAAGCCAGGCTGCGCGAGATCCGGACAACT 1433
DB 1322 GGCAGAGGTGCAAGGTTCAAGAGAGAG-----AAGCT 1354
QY 1434 AAGATTTATGCTCTACTGACAAACCAACCAACCACTAGCTGCTGCTCATTTAC 1493
DB 1355 TCGAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1414
QY 1494 ACTTTTATCATCACTTGTGATGATCAGAGAAAGAAATCAAGCTGCTGGACTCTCAG 1553
DB 1415 TCTGTATGCAATAAACCCTCATATACGTCACCAAGTCTTGGCTTACCTTATCTTTTC 1474
QY 1554 AGACAGCTGTTTACAGTACGTCGCTGAGCCCTTATGAGGAGAGGAGGCTTAAGTCCA 1613
DB 1475 TAACAGAGAGTATTAATTAATCTTTTACAGCTTTGAGAGCTCTGATTAATCTTCCA 1534
QY 1614 GTCAGTCAACTGATAGGAGAGCCCTTATGATGATGATGAGAGAGAGCTCTCCAGAAAT 1673
DB 1535 ATCTGTCAACTCATGATGATTAATCTTAAGATGATGATTAATCACTTGCACCTTT 1594
QY 1674 GAACCCCCCGCTTGGGGCGGCGAGATTTGTCATCTCTCCAGTACCATGAGCTT 1733
DB 1595 AATGAAAAACCTCTCGGCGAGAGATTCATGAGGCTTGCAGCTTTCTCATATGATTT 1654
QY 1734 TTATGAGTCAAGATGTCATGCTTTGGCTGCTCCGCTACCGATTA 1779
DB 1655 TTTTGATTAAGAAATGCCAAAGTGTGCTGTGATCATGTAATAA 1700

```

RESULT 10  
 US-09-635-923-1  
 ; Sequence 1, Application US/09635923  
 ; Patent No. 6426209  
 GENERAL INFORMATION:  
 APPLICANT: Maty Ayal-Herskowitz et al.  
 TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
 EXPRESSING RECOMBINANT HEPARANASE  
 AND METHODS OF PURIFYING SAME  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Mark M. Friedmann c/o Anthony Castorina  
 STREET: 2001 Jefferson Davis Highway, Suite 207  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: United States of America  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 COMPUTER: PC

```

OPERATING SYSTEM: MS DOS version 6.2,
SOFTWARE: Word for Windows version 3.11
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/635,923
  FILING DATE: 10-Aug-2000
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/09/487,716
  FILING DATE: 19-Jan-2000
  APPLICATION NUMBER: 09/071,618
  FILING DATE: May 1, 1998
  APPLICATION NUMBER: 09/071,739
  FILING DATE: May 1, 1998
  APPLICATION NUMBER: 08/922,180
  FILING DATE: September 2, 1997
ATTORNEY/AGENT INFORMATION:
  NAME: Friedmann, Mark M.
  REGISTRATION NUMBER: 33,883
  REFERENCE/DOCKET NUMBER: 910/16
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 972-3-5625554
  TELEFAX: 972-3-5625554
  TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1721
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-635-923-1

Query Match      15.7%; Score 278.8; DB 4; Length 1721;
Best Local Similarity 52.4%; Pred. No. 3.8e-76;
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

```

1014 TGATGCCGGGTGTGCTAAGGTATGACATTCCTGAAAACCTGGCTGTATGACACTCTC 1073
1015 GATATGACGACGACTGCTACACGAGGAAATTTTCTAAACCTGATGTATGACATTTTAT 1021
1016 TGACCAATATGAGAAATTCAGAAAGGTATTAATACATCACTCCGAGAAAGATTG 1133
1017 TTAATCTGTCAAAAATTTTCCAGGTGTGAGACACAGGCTCGGCAAGAGTCTG 1081
1018 GCTTGAAGGTGTGTGACCACTCACTGAGGCAACAAATCTATCCGATTCCTATGC 1193
1019 GTTATGAGAAACCAAGCTCTCATATGAGGCGAGGCGCTTCTATCCGACACTTTCG 1141
1020 TGAGAGATTTCTTATGTGTGAACCTTTAGAAAGCTGTGCGCAATCAGAGGATGTGCT 1253
1021 AGCTGGCTTTATGTGTGTGATTAATTGGGCTGTGAGGCCGATGGAATGAGATGAT 1201
1022 GATACGCACTCATTTTGTGACATGATGATCAATCACTGTGTGACAGAAATTTTAA 1313
1023 GATGAGCAAGATATTTCTTGTGAGAGAAATCACTATTAATGATGAGAAATTTGATCC 1261
1024 ATTTCCAGACTACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1373
1025 TTAACTGATTTATGTGTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
1026 TGTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1433
1027 GGCAAGGTGTCAAGGTGTCAAGGAGG-----AAGCT 1354
1028 AAGATTTATGTCTCACTGTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1493
1029 TCGATGTACCTTCATTTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1414
1030 ACTTTTATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
1031 TCTGTATGCAATTAACCTCATATGATGATGATGATGATGATGATGATGATGATGATG 1474
1032 AGACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1613
1033 TTAACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1534
1034 GTGAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1673
1035 ATCTGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1594
1036 GAAACCCCGCCCTTGTGAGGCGGCGGCAATGATGATGATGATGATGATGATGATGAT 1733
1037 AATGAAACCACTCTCGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1654
1038 TTAATGTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
1039 TTTGTATTAAGAAATGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1700

RESULT 11
US-09-635-923-3
Sequence 3, Application US/09635923
Patent No. 6426209
GENERAL INFORMATION:
APPLICANT: Mary Ayal-Herszkovitz et al.
TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR
EXPRESSING RECOMBINANT HEPARANASE
AND METHODS OF PURIFYING SAME
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mark M. Friedmann c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: PC

```





Fri May 14 13:01:53 2004

us-10-088-676-1.rn1

Page 12

OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted to an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/487,716A  
FILING DATE: 19-Jan-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/071,618  
FILING DATE: May 1, 1998  
APPLICATION NUMBER: 09/071,739  
FILING DATE: May 1, 1998  
APPLICATION NUMBER: 08/922,180  
FILING DATE: September 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-487-716A-1  
Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;  
QY 474 AGATTAACGAAAGGCTGCAAGATTGCGGACGACCTGATGTTATGCTGCTGCTCAAG 533  
DB 425 AGTAAACGAGATATTGGCAAAATATGATCCATCCCTGATGAGGAGAACTTACG 484  
QY 534 GAGAAAGGACCTGATGATGATGCTGCTTCTTAAAGAGCAATCTCAATATTACAG 593  
DB 485 GTTGAGATGGCCCTACAGAGGCAATTGCTACCTCCAGAACACTACGAAAGAAATTGA 544  
QY 594 TAATCTGATTAACAGCCAGGCTCTCAAGCAAACTTTAACTTCTGATGCTGCTG 653  
DB 545 GAAAGACCTACTCAAGAAAGCTCTGATGATGCTATACCTTTGCAATGCTCAG 604  
QY 654 ACTCAGCTGATATTTCTTAATGCACTGCTGCTAATCCCAATACCTCTGAGAG 713  
DB 605 ACTGAGCTGATCTTGGCTTAATGCTTAAAGAGGCTTAAAGACAGAGATTGCACTGAG 664  
QY 714 TTCTAGTCCCTGAGTCTTGAAGTACAGGCGGACGCAAAAGTACACATTCTTGGGA 773  
DB 665 TTCTAAAGCTCAGTGTCTGCTGACTAATGCTCTTCCAAAGGGGTAAACATTTCTTGGGA 724  
QY 774 ACTGGTAAATAGCCCAATTAATGCAATGCAATGCAATGCGGAGAGTAAATGAGAGCA 833  
DB 725 ACTAGGCAATGAACTTAACAGTTCTTAAAGAGGCTGATATTTTCAATGAGGTGCA 784  
QY 834 GTTGGGAAAGATTAATCACTGAGTGAAGGCTGTTGCAAGCCCAATCCGATTTATCCAG 893  
DB 785 GTTGGGAAAGATTAATCACTGAGTGAAGGCTGTTGCAAGCCCAATCCGATTTATCCAG 841  
QY 894 AGCAGCTTAATGAGCCCTTAATATTTGGGCGGCGGAGAGAAATGTCATGCGCTCTAGA 953  
DB 842 TGCAAAATCTATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
QY 954 TGATTCATGAGGCTGAGCAAGAGTACAGATGAGTGAATGAGTGAATGAGTGAATGAGT 1013  
DB 902 GAGCTTCTGAGAGGCTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 961  
QY 1014 TGATGCGGCTGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1073

DB 962 GATGAGGAGCTGCTACAGGAGATTTCTTAAACCTGATGATGATGATGATGATGAT 1021  
QY 1074 TACCAAGATTGAGAAATTCAGAAAGTGTAAATTAATCACTCCAGAGAAAGATTTG 1133  
DB 1022 TTCACTGTGAGAAAGATTTCCAGGCTGTTGAGAGACAGGCTGCGCAAGAGCTG 1081  
QY 1134 GCTTGAAGTGTGATGACCACTGAGTGAAGGCAACAACATCTATCCGATTTCTATG 1193  
DB 1082 GTTAAAGAAACAGCTCTGATATGAGAGGCGGCTGCTGATCCGACCTTTG 1141  
QY 1194 TCCAGATTTCTTAAAGTGAACACTTAAAGATGCTGCGCAATCAGGAGTATGCTG 1253  
DB 1142 AGCTGCTTTATGCTGCTGATTAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201  
QY 1254 GATAGGAGCTGATTTTGAACATGATGATGATGATGATGATGATGATGATGATGAT 1313  
DB 1202 GATGAGGAGATTTCTTGAAGAGAGAACTACATTTTATGATGAGAAACTTGCATCC 1261  
QY 1314 ATTACAGACTAATGCTCTCTCTCTCTCAAGGCTGATGAGGCGCCAAAGCTTGGC 1373  
DB 1262 TTACCTGATTTATGCTATCTCTCTCTCTCAAGAAATGATGAGGACCAAGCTGTAAT 1321  
QY 1374 TGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1433  
DB 1322 GGCAGGCTGAGAGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1354  
QY 1434 AAGATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1493  
DB 1355 TCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1414  
QY 1494 ACTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1553  
DB 1415 TCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1474  
QY 1554 AAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1613  
DB 1475 TAAACAGAGAGATGATTAATCTCTTAAAGCTTTGAGGAGCTGATGATTTTCCAA 1534  
QY 1614 GTCACTGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1673  
DB 1535 ATCTGCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594  
QY 1674 GAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1733  
DB 1595 AATGAGAAACCTCTGCGGCGGAGAGATTCATGAGGCTTCTCATATAGTTT 1654  
QY 1734 TTATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1779  
DB 1655 TTTTGTATTAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700  
RESULT 13  
US-09-487-716A-3  
Sequence 3, Application US/09487716A  
Patent No. 6475763  
GENERAL INFORMATION:  
APPLICANT: Maty Ayal-Hershkovitz et al.  
TITLE OF INVENTION: GENETICALLY MODIFIED CELLS AND METHODS FOR  
EXPRESSING RECOMBINANT HEPARANASE  
AND METHODS OF PURIFYING SAME  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Mark M. Friedman c/o Anthony Castorina  
STREET: 2001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: PC  
OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11

```

SOFTWARE: Word for Windows version 2.0 converted to an ASCII file
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/487,716A
FILING DATE: 19-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,618
FILING DATE: May 1, 1998
APPLICATION NUMBER: 09/071,739
FILING DATE: May 1, 1998
APPLICATION NUMBER: 08/922,180
FILING DATE: September 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 910/16
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-562553
TELEFAX: 972-3-562554
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1721
TYPE: nucleic acid
STRANDEDNESS: Double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-487-716A-3

Query Match      15.7%   Score 278.8; DB 4; Length 1721;
Best Local Similarity 52.4%; Pred. No. 3.8e-76;
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

```

```

1074 TGACCAAGATTAGAAAATTCAGAAAGTGTAAATACATACCTCCAGGAAGAAAGATTG 1133
1022 TTCATCTGTGCAAAAAGTTTCCAGGTGTGAGGACACGAGCCCTGGCAAGAGTCTG 1081
1134 GCTTGAAGGTGTGTGACCACTCAGCTGAGGACCAAAACATCTATCCGATTCCTATGC 1193
1082 GTTAGAGAAACAAAGCTGTGATATGAGGGGAGGCGCCCTTGCTGACACCTTTGC 1141
1194 TCGAGATCTTATGTTGTTGAACACTTTAGATGTGCGCAATCGAGCATGTATGTCTG 1253
1142 AGCTGCTTATGTGTGCTGATTAATGGGCTGTACGCCGGAATGGAAATGAAAGTGT 1201
1254 GATAGGCACTGATTTTGTGACCATGATTAATACCTCTGTGACCAATTTTAACC 1313
1202 GATGAGGCAAGATATCTTTGAGCAGGAACATACCATTTAATGTGATGAAAACCTGATCC 1261
1314 ATTACAGACTACTGAGCTCTCTCTCTACAGGCGCTGATCGGCCCAAGCTTTGCG 1373
1262 TTACCTGATTTATGTGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
1374 TGTGATGTGCTGTGGCTCCAGCGGAGCCACGGCTTGGCCGAGTATCCGGACAAACT 1433
1322 GCAAGCGTCAAGGTTCAAGAGAGAG-----AGCT 1354
1434 AAGGATTTATGCTCATCTGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1493
1355 TCGAGTATCTTCACTTATGACAAACCACTGACATCAAGTATTAAGAGAGATTTAAC 1414
1494 ACTTTTATCATCAACTGATCATGATCAAGAAAGAAATCAAGCTGTGGGACTCTGAG 1553
1415 TGTGATGCAATTAACCTTCATTAACGATCAACCAAGTATGAGGTATCCCTATCTTTTC 1474
1554 AAGCAAGCTGTGTCACGATACCTGTGACGCGCTTATGAGGAGAGGCGCTTAAGTCAA 1613
1475 TAACAAGCAAGGTATTAATACCTTTTAAGACCTTTGAGCTTCAAGATTAATCTTCAA 1534
1614 GTCAGTCACTGATGAGCGACGCTTATGATGTGTGACAGGAGACCTTCCAGAAAT 1673
1535 ATCTGCACTCATGATGCTTATCTTAAGATGATGATGATTAAGATGATGATGATGATG 1594
1674 GAAAGCGCGCGCGCTTGGGCGCGCGGACATGTGATCTCTCCAGTCAACATGAGGCTT 1733
1595 AATGAAAAAATCTTCCGCGCGGAGGATCACTGGGCTTGGCAGGCTTCTCATATATATTT 1654
1734 TTATGTCATGAATGTCATGCTTTGGGCTGCGGCTAACGATTA 1779
1655 TTTTGTATGAAGAAATGCCAAGTTGCTGCTTGCATCTGAATATA 1700

RESULT 14
US-09-322-977-1
Sequence 1, Application US/09322977.
Patent No. 6531129
GENERAL INFORMATION:
APPLICANT: ILS Becker et al.
TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES
TITLE OF INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Twinhead Slimnote-890TX
OPERATING SYSTEM: MS DOS version 6.2,
SOFTWARE: Word for Windows version 3.11
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/322,977  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/922,180  
 FILING DATE: September 2, 1997  
 APPLICATION NUMBER: 09/071,739  
 FILING DATE: May 1, 1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Friedman, Mark M.  
 REGISTRATION NUMBER: 33,883  
 REFERENCE/DOCKET NUMBER: 910/21  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 972-3-562553  
 TELEFAX: 972-3-562554  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1721  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-09-322-977-1

Query Match 15.7%; Score 278.8; DB 4; Length 1721;  
 Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
 Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

474 AGATAACAAAGAGGCTGCAAGATGCGCCAGACCCCTGATGATGCTGCTCCAAAG 533  
 DB 425 AGTAAACAGAGATATTTGCAATATGATCATCTCTCTGATGATGAGAGAGTTACG 484  
 QY 534 GAGAGAGCAGCTCAGATGATGATGCTGCTCTTCTTAAAGAGCAATTCCTCAATCTTACG 593  
 DB 485 GTTGAGATGGCCCTACACAGAGCAATGCTACTCCGAAACATCACAGAAAGTTCAA 544  
 QY 594 TTAATCTATATTAACACCGAGCTCTTACAAACCTTATATCTTCTGATGCTCTG 653  
 DB 545 GAAACAGACCTACTCAAGAAAGCTGTGATGATGCTTACCTTTGCAAACTGCTCAGG 604  
 QY 654 ACTGACCTGATATTTGCTCTTAAATGCACTGCTGATATCCCAATTAATCTCGAAACAG 713  
 DB 605 ACTGATCTTGAATCTTTGGCCCTAAATGCGTTATTAAGAACAGAGATTTGCAATGAAACAG 664  
 QY 714 TTCTAGTCCCTGATGCTGTTGAAATGACGCGCCAGCAAAAGTACAACTTTCTGGGA 773  
 DB 665 TTCTAAATGCTCAAGTTGCTCTGCTGCTACTGCTCTTCCAAAGGATTAATCAATCTTCTGGA 724  
 QY 774 ACTGGTAAATAGGCCAAATATCTATCGGACCATGATGCGCGGAGTAAATGCGACCCA 833  
 DB 725 ACTAGGCAATGAACCTTAACAGTTTCTTAAAGAGCTGATATTTTATCAATGGGTGCA 784  
 QY 834 GTTGGGAAAGATTAATCACTGCTGAAAGAGCTTTGACGCCATCCGGATTAATTCAG 893  
 DB 785 GTTAGGAGAAATATATTAATGCAATAAC--TTCTAAGAAAGTCCACCTTCAAAA 841  
 QY 894 AGCCAGCTTATATGCGCTTATATTTGGCGCGCGAGAGAAAGATTCATCCCTCTCAGA 953  
 DB 842 TGCAAAACCTATATGCTCTGATGTTGTCAGCTCCGAAAGAAAGCGCTAAGATGCTGAA 901  
 QY 954 TGAATTCATGAAGGTGCGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1013  
 DB 902 GAGCTTCTCTGAAGGCTGTGTGAGAGAGTATGATGATGATGATGATGATGATGATGAT 961  
 QY 1014 TGATGCGCGGCTGCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
 DB 962 GATGAGAGAGCTGCTACAGAGAGATTTTCTAAACCTGATGATGATGATGATGATGAT 1021  
 QY 1074 TGACCAATTAAGAAATTAAGAAAGTGTATTAATCACTCCAGAAAGAAAGATTG 1133  
 DB 1022 TTCACTGTGCAAAAGTTTTCAGGTGTGAGAGAGCAAGGCTGCAAGAGGCTG 1081  
 QY 1134 GCTTGAAGGTGTGAGAGCAAGCTGAGTGAAGAGCAAAAGATTAATCCGATTCATGCT 1193

DB 1082 GTTAGGAGAAACAGCTCTGATATGAGAGGAGGCGCCCTGCTATCCAGACCTTTGC 1141  
 QY 1194 TGCAGATTTCTTATGTTGAACACTTTAGAAATGCTGCCAATCAGGCGATTAATGCT 1253  
 DB 1142 AGCTGGCTTATATGCTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1201  
 QY 1254 GATACGCACTATTTTTTGGACATGATTAATCACTGCTGATGATGATGATGATGATGAT 1313  
 DB 1202 GATAGAGGATATTTCTTTGAGAGAGAACTACATTTATGATGATGATGATGATGATGAT 1261  
 QY 1314 ATTAACAGACTACTGCT 1373  
 DB 1262 TTATCTGATATTTGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321  
 QY 1374 TGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1433  
 DB 1322 GGCAGGCTGCAAGGTTCAAGAGAGG-----AGCT 1354  
 QY 1434 AAGATTTATGCTCACTGCAACAAACCAACCAACCAACCAACCAACCAACCAACCAAC 1493  
 DB 1355 TCGATATACCTTCAATGCAAAACATCAATCCAAAGGATTAAGAGAGATTTAC 1414  
 QY 1494 ACTTTTATCAATCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1553  
 DB 1415 TGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1474  
 QY 1554 AGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613  
 DB 1475 TAAAGAGAGAGGATTAATTAATCTTCTTAAAGCTTTGGAGCTGATGATTAATTTCAA 1534  
 QY 1614 GTCATGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673  
 DB 1535 ATCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594  
 QY 1674 GAAGCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1733  
 DB 1595 AATGAGAAACCTCTCGGCGCAGAGATGATGATGATGATGATGATGATGATGATGAT 1654  
 QY 1734 TTATGCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 DB 1655 TTTTGTGATGAAGAAATGCAAGTGTGCTGCTGATGATGATGATGATGATGATGAT 1700

RESULT 15  
 US-09-322-977-3  
 ; Sequence 3, Application US/09322977  
 ; Patent No. 6531129  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Irls Pecker et al.  
 ; TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
 ; TITLE OR INVENTION: AND THEIR USE IN RESEARCH AND MEDICAL  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
 ; STREET: 2001 Jefferson Davis Highway, Suite 207  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: United States of America  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
 ; COMPUTER: Twinhead\* Slimnote-8907X  
 ; OPERATING SYSTEM: MS DOS version 6.2,  
 ; SOFTWARE: Word for Windows version 3.11  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/322,977  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/922,180

Fri May 14 13:01:53 2004

us-10-088-676-1.rn1

Page 15

FILING DATE: September 2, 1997  
APPLICATION NUMBER: 09/071,739  
FILING DATE: May 1, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fieldman, Mark W.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 910/21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1721  
TYPE: Nucleic acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
US-09-322-977-3

Query Match 15.7% Score 278.8; DB 4; Length 1721;  
Best Local Similarity 52.4%; Pred. No. 3.8e-76;  
Matches 684; Conservative 0; Mismatches 592; Indels 30; Gaps 2;

QY 474 AGATTAAGAAAGCTGCAAGATTGCCAGCAGCCGATGTTATGCTGCTCCAAAG 533  
DB 425 AGTCAACAGAGATTTGCAAAATGATCCATCCCTCTGATGTGAGAGAAATTACG 484  
QY 534 GAGAGAGCAGCTCAGATGATCTGTTCTTAAAGAGCAATTCCTCACTTAC 593  
DB 485 GTTGAATGCGCCTACAGAGCAATGCTACTCCGAAACATACAGAAAAGTTCAA 544  
QY 584 TAATCTATTAACAGCAGCTCTCTGACAACTTTATACCTTTGCTGATTCCTCG 653  
DB 545 GAAAGCAGCTACTAGAGAGCTGTAGATGTGTATACCTTTGCAACCTGCTCAGG 604  
QY 654 ACTCAGCTGATATTGCTCTAAATGCACTGCTGTATCCCAATTACTCTGAAACAG 713  
DB 605 ACTGAGCTGATCTTTGGCTTAATGCGTTATTAAGAACAGCAATTTGCAAGTGAACAG 664  
QY 714 TTCTAGTCCCTGATCTGTGTAAGTACAGCCGCAAGAAAAGTCAACATTTCTTGGGA 773  
DB 665 TTCTATATGCTCAGTGTCTCTGAGCTACTGCTCTTCCAGGAGTATTAACATTTCTTGGGA 724  
QY 774 ACTGGATATGAGCCAAATTAATCGAGCATGATGCGGAGAGTAAATGCGACGCA 833  
DB 725 ACTAGGCAATGATCACTTAAGTCTTAAAGAGCTGATTTTCAATGATGGTGCA 784  
QY 834 GTTGGAAAGATTACATCCAGCTGAGAGCCTGTGCAAGCCCATCGGATTTATCCAG 893  
DB 785 GTTAGAGAGATATATTCATTCATTAAC---TTCTAAGAAAAGTCCACCTTCMAAAA 841  
QY 894 AGCCAGCTTATATGGCCCTAATATTTGGGGGCGGAGAGAAATGTCATCGCCCTCTAGA 953  
DB 842 TGCATAAATCTATGATGCTGATGTTGTCAGCTCTGAAAGAAAGAGCTGAAGTGTGAA 901  
QY 954 TGGATTCAATGAAGTGGAGAGAGTACAGTATGCACTTACCTGGCAACATTTGCTACAT 1013  
DB 902 GAGCTTCTGAAGGCTGGTGGAGAGTATGATTCAGTTACATGGCATCACTACTAATTT 961  
QY 1014 TGATGGCCGGGTGTGTCAGAGTGTATGAGTCTCTGAAAACCTGCGCTGTAGACACTCTC 1073  
DB 962 GATAGGACGAGCTGTACAGAGAGAGATTTCTTAACCTGATGTATTTGAACTTTTAT 1021  
QY 1074 TGACCAATTAAGAAATTCAGAAAAGTGTATATACATACCTCCAGAGAAAAGATTTG 1133  
DB 1022 TTCACTGTGCAAAAAGTTTCCAGGTGTGAGAGCACAGGCTGTGCAAGAAAGTCTG 1081  
QY 1134 GCTTAAGGTGTGAGCACTCAGCTGAGAGCAAAACATCTATCCGATTCCTATGC 1193  
DB 1082 GTTAGAGAAACAGCTGTGATATGAGAGGCGAGCGCCCTGTATCCGACCTTTGCG 1141  
QY 1194 TGCAGATTCTTATGTTGAACATTTAGATGCTGGCAATCAGGGCATTTGATGTGT 1253

DB 1142 AGCTGCTTATGCTGAGATTAATTGGGCTGTCAAGCCGCAATGGAAATGAAGTGT 1201  
QY 1254 GATAGGCACTCATTTTGTGACATGATATACATCACTCTGTGACAGAAATTTTAAACC 1313  
DB 1202 GATAGGCAAGATTTCTTTGAGCAGGAAATCACTATTAATGATGAAACTTGATCC 1261  
QY 1314 ATTACCAATCACTGCTCTCTCTCTCTCAAGGCGCTGATGCGCCCAAAAGTTTGGC 1373  
DB 1262 TTACTCTGATTTATGCTATCTCTCTCTCAAGAAATTTGGTGGGCAACAAAGTGTAT 1321  
QY 1374 TGTGATGTGCTGGGCTCCAGGAGAGCAAGCGCTGCGCAATGATCCGGACAAC 1433  
DB 1322 GGCAGCGTGAAGTTCAAGAGAGAG-----AAGCT 1354  
QY 1434 AAGATTTATGCTCACTGCAACCAACCAACCAACCACTACGTTGCTGCTCATTAAC 1493  
DB 1355 TCGAGTATACCTTCAATGCAACCACTGACCAATCAAGTATTAAGAGAGATTTAAC 1414  
QY 1494 ACTTTTATCACTCACTGATGATGATCAAGAAAGAAATCAAGCTGCTGGGACTCTGAG 1553  
DB 1415 TCTGATGCAATTAACCTTCATTAAGTACCAAGTACTTGGGCTTACCTATCTTTTTC 1474  
QY 1554 AGCAAGCTGTTCAACAGTACTGCTGCAAGCTTATGAGGAGAGGCTTAAAGTCAA 1613  
DB 1475 TAAACAGAGAGTGAATTAATACCTTCTAAGACCTTTGGGACTCATGATTTCTTCAA 1534  
QY 1614 GTCACTGCACTGATGCTGCAAGCTTATGATGATGATGATGATGATGATGATGATGAT 1673  
DB 1535 ATCTGTCAACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594  
QY 1674 GAAGCCCGCCCGCTTGGGCGGCGGAGCATTTGATCACTCCCTGATCAGCATGAGGCTT 1733  
DB 1595 AATGAAAAACCTTCGCGGCGAGAAATTCATGAGGCTTGCAGGCTTCTCATATAGTTT 1654  
QY 1734 TTATGTGTCAGAAATGTCATGCTTTGGCTGCGGCTTACCATTA 1779  
DB 1655 TTTTGTATGAAGAAATGCAAGTGTGCTGCTGATCTGAAATATA 1700

Search completed: May 13, 2004, 15:40:01  
Job time: 147 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:34:34 ; Search time 6955 Seconds

(without alignments)  
11086.593 Million cell updates/sec

Title: US-10-088-676-1

Perfect score: 1779

Sequence: 1 atgagggtcgttctgtcctc.....tggcctgcgcgtacgcgataa 1779

Scoring table: IDENTITY NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: 1: gb\_ba: 2: gb\_bhg: 3: gb\_in: 4: gb\_ov: 5: gb\_ov: 6: gb\_ov: 7: gb\_ov: 8: gb\_ov: 9: gb\_ov: 10: gb\_ov: 11: gb\_ov: 12: gb\_ov: 13: gb\_ov: 14: gb\_ov: 15: gb\_ov: 16: gb\_ov: 17: gb\_ov: 18: gb\_ov: 19: gb\_ov: 20: gb\_ov: 21: gb\_ov: 22: gb\_ov: 23: gb\_ov: 24: gb\_ov: 25: gb\_ov: 26: gb\_ov: 27: gb\_ov: 28: gb\_ov: 29: gb\_ov: 30: gb\_ov: 31: gb\_ov: 32: gb\_ov: 33: gb\_ov: 34: gb\_ov: 35: gb\_ov: 36: gb\_ov: 37: gb\_ov: 38: gb\_ov: 39: gb\_ov: 40: gb\_ov: 41: gb\_ov:

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1779	100.0	1779	6	AX100935 Sequence
2	1774.2	99.7	2096	6	AF282887 Homo sapi
3	1774.2	99.7	2353	6	AX299719 Homo sapi
4	1773.8	99.7	2636	6	AX180830 Sequence
5	1769	99.4	1847	6	AX393820 Sequence
6	1694.2	95.2	1972	6	HS429720 Homo sapi
7	1441.4	81.0	1685	6	AX393824 Sequence
8	1417.8	79.7	1922	6	AF282886 Homo sapi
9	1417.8	79.7	2326	6	AX286343 Sequence
10	1417.8	79.7	2462	6	AX180832 Sequence
11	1411.4	79.3	2064	9	AX094037 Homo sapi
12	1411	79.3	1673	6	AX393822 Sequence
13	1289.8	72.5	3943	6	AX190449 Sequence
14	1132.2	63.6	2369	6	AX280370 Sequence
15	1095.4	61.6	1760	9	AF282885 Homo sapi
16	1095.4	61.6	2300	6	AX180834 Sequence
17	1093.4	61.5	1511	6	AX393826 Sequence
18	432.4	24.3	468	6	AX259914 Sequence
19	426.4	24.0	607	6	AX280372 Sequence
20	375.8	21.1	400	6	AX180851 Sequence
21	292.8	16.5	1605	5	AY037007 Gallus ga
22	290.4	16.3	123482	9	AL445251 Human DNA
23	290.4	16.3	172219	2	AL442004 Homo sapi
24	289.2	16.3	1662	4	AF281160 Bos tauru
25	285.2	16.0	1736	10	AY077467 Mus muscu
26	283.2	16.0	1912	10	AF359507 Mus muscu
27	279.4	15.7	1754	10	AF359508 Rattus no
28	278.8	15.7	1721	6	AR080679 Sequence
29	278.8	15.7	1721	6	AR080680 Sequence
30	278.8	15.7	1721	6	AR125603 Sequence
31	278.8	15.7	1721	6	AR125604 Sequence
32	278.8	15.7	1721	6	AR194189 Sequence
33	278.8	15.7	1721	6	AR194190 Sequence
34	278.8	15.7	1721	6	AR221285 Sequence
35	278.8	15.7	1721	6	AR221286 Sequence
36	278.8	15.7	1721	6	AR243303 Sequence
37	278.8	15.7	1721	6	AR243304 Sequence
38	278.8	15.7	1721	6	AR287435 Sequence
39	278.8	15.7	1721	6	AR287436 Sequence
40	278.8	15.7	1721	6	BD074427 Polynucle
41	278.8	15.7	1721	6	BD074428 Polynucle
42	278.8	15.7	1721	6	BD193336 Hepatarnas
43	278.8	15.7	1721	6	BD193337 Hepatarnas
44	278.8	15.7	1721	6	BD205238 Cells gen
45	278.8	15.7	1721	6	BD205239 Cells gen

# ALIGNMENTS

RESULT 1  
AX100935  
LOCUS AX100935 1779 bp DNA PAT 11-APR-2001  
DEFINITION Sequence 1 from Patent WO0121814.  
ACCESSION AX100935  
VERSION AX100935.1 GI:13619821  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Duecker, K. and Sirenberg, C.  
Hepatarnase-2, a member of the hepatarnase protein family  
Patent: WO 0121814-A 1 29-MAR-2001;  
JOURNAL

FEATURES  
MERCK PATENT GmbH (DE)

Location/Qualifiers  
Source

1..1779  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1..1779  
/note="unassigned protein product"

/codon\_start=1  
/protein\_id="CAC36496.1"  
/db\_xref="GI:13619822"  
/translation="MRVLCAPFAPMSNSRPPACPAALYALILHLISLSQAGDR  
RPLPVRAAGLKEKTLILDVSTKRNPRVTNENPLSLDPSIHDGWLDEFSLRLV  
TLARGLSPALFRGGRPTLPRONLRNPAKSGRPYIKNVEDIVRSDVLDK  
QKCKTIAQHPDVLVLOREKAQPHLVLEQPSNTSYMLITPAKSLKLPADCSG  
LHLIPALNALRNPNSWNSLSLILKTSASRKNVILSDGPNPNTWTHGRVAVG  
SOLKDYIOLKSLIOPRIYRSALYPIGRKNVIALDGFNVAISTVDATWQ  
HCYIDGRVAVKMDPLKTRLDLSDQIRIKOVNTYTGKILWEGVVTAGGTNN  
LSDSYAGFLMNTGLANOGIDVIRHSPFDHGYNHLVDNPNLIPYMLSLYKR  
LIGPKVLAVHAGLQKRPGRGIRDKLFIYACTNHNHNVDNPSITLFIHLRSR  
KRIKAGTLADKLHOYLLOPYGSGELKSKSVLNGQLVMTDDGLPELKRPLRAG  
RIVIPPTMGTFYVKNVNAACRYR"

ORIGIN

Query Match 100.0%; Score 1779; DB 6; Length 1779;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAGAGGTGCTTTTGCTTCCCTGGAAGCCATGCTCCAGACACTCCGCCCCCGCG 60  
DB 1 ATAGAGGTGCTTTTGCTTCCCTGGAAGCCATGCTCCAGACACTCCGCCCCCGCG 60  
QY 61 TGCTTAGCCCCGGGGGCTCTACTTGCTGCTGCTGCTCCTCCATCTCCCTTCTCCAG 120  
DB 61 TGCTTAGCCCCGGGGGCTCTACTTGCTGCTGCTGCTCCTCCATCTCCCTTCTCCAG 120  
QY 121 GCTGAGACAGAGACCCCTTGCTGACAGAGCTGACAGTTTGAAGAAAAGACCTTG 180  
DB 121 GCTGAGACAGAGACCCCTTGCTGACAGAGCTGACAGTTTGAAGAAAAGACCTTG 180  
QY 121 GCTGAGACAGAGACCCCTTGCTGACAGAGCTGACAGTTTGAAGAAAAGACCTTG 180  
DB 121 GCTGAGACAGAGACCCCTTGCTGACAGAGCTGACAGTTTGAAGAAAAGACCTTG 180  
QY 181 ATTCTACTGATGAGACCAAGAAACCCAGTCAAGACATGATGAACTTCTCTCT 240  
DB 181 ATTCTACTGATGAGACCAAGAAACCCAGTCAAGACATGATGAACTTCTCTCT 240  
QY 241 CTGCAAGTGAATCGTTCATCATCATGATGCTGCTGCTGCTTCTTAAGCTCAAGCGC 300  
DB 241 CTGCAAGTGAATCGTTCATCATCATGATGCTGCTGCTGCTTCTTAAGCTCAAGCGC 300  
QY 301 TTGCTGACCCCTGAGCCGGGGACCTTGGCCGCTTCTGCGCTGCGGGGGCAAAAGAC 360  
DB 301 TTGCTGACCCCTGAGCCGGGGACCTTGGCCGCTTCTGCGCTGCGGGGGCAAAAGAC 360  
QY 361 GACTTCTGAGTTTCAAGACCTGAGAAACCCGCGCAAAAGCCGCGGGGCCCGCG 420  
DB 361 GACTTCTGAGTTTCAAGACCTGAGAAACCCGCGCAAAAGCCGCGGGGCCCGCG 420  
QY 421 GATTACTGATCTCAAAACCTATGAGAGATGATTTGCTGAGAGTGTCCCTTAAGTAA 480  
DB 421 GATTACTGATCTCAAAACCTATGAGAGATGATTTGCTGAGAGTGTCCCTTAAGTAA 480  
QY 481 CAGAAAGGCTGCAAGATTTGCCAGACCTGATGATGCTGCTGCTCAAAAGGAGAG 540  
DB 481 CAGAAAGGCTGCAAGATTTGCCAGACCTGATGATGCTGCTGCTCAAAAGGAGAG 540  
QY 541 GCGCTCAATGATCTGCTTCTTAAGAGAGCAATTTCCATCTTAAGTATCTC 600  
DB 541 GCGCTCAATGATCTGCTTCTTAAGAGAGCAATTTCCATCTTAAGTATCTC 600  
QY 601 ATATTAAAGCAGCTCTAGCAAACTTTAACTTTGATGCTGCTGAGCTCCAC 660  
DB 601 ATATTAAAGCAGCTCTAGCAAACTTTAACTTTGATGCTGCTGAGCTCCAC 660

QY 661 CTGATATTGCTCTAAATGCACTGCGTGTAAATCCCAATACTCCTGAAACAGTTCTAGT 720  
DB 661 CTGATATTGCTCTAAATGCACTGCGTGTAAATCCCAATACTCCTGAAACAGTTCTAGT 720  
QY 721 GCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAGTAAACATTTTGGGAACTGGGT 780  
DB 721 GCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAGTAAACATTTTGGGAACTGGGT 780  
QY 781 AATGAGCCAAATACTATCGGACCATGCGCGGCACTTAATGACAGCCAGTTGGGA 840  
DB 781 AATGAGCCAAATACTATCGGACCATGCGCGGCACTTAATGACAGCCAGTTGGGA 840  
QY 841 AAGGATTAATCCGCTGGAAGAGCTTTCAGGCCATCCGGAATTTTCCAGAGCCAGC 900  
DB 841 AAGGATTAATCCGCTGGAAGAGCTTTCAGGCCATCCGGAATTTTCCAGAGCCAGC 900  
QY 901 TTATATGCCCCCTAAATTTGGGCGCGCCAGGAAAGATGTCAATCCGCTCTAGATGATTC 960  
DB 901 TTATATGCCCCCTAAATTTGGGCGCGCCAGGAAAGATGTCAATCCGCTCTAGATGATTC 960  
QY 961 ATGAAAGTGGGAGAGAACTACATGATGCACTTCTGCAACATTTGCTATGATGCTC 1020  
DB 961 ATGAAAGTGGGAGAGAACTACATGATGCACTTCTGCAACATTTGCTATGATGCTC 1020  
QY 1021 CGGCTGCTCAAGGTGATGCACTTCTGCAAACTGCTGTAGACACATCTCTGACACAG 1080  
DB 1021 CGGCTGCTCAAGGTGATGCACTTCTGCAAACTGCTGTAGACACATCTCTGACACAG 1080  
QY 1081 ATTGAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
DB 1081 ATTGAGAAATTCAGAAAGTGTATATATATATATATATATATATATATATATATAT 1140  
QY 1141 GGTGTGTGACACACTGAGCTGAGGACAAACATATCCGATTCCTATGCTGACAGGA 1200  
DB 1141 GGTGTGTGACACACTGAGCTGAGGACAAACATATATATATATATATATATATATAT 1200  
QY 1201 TTCTTATGTTGAACACTTTAGAAATGCTGCGCAATGATGCTGATGCTGATGCTG 1260  
DB 1201 TTCTTATGTTGAACACTTTAGAAATGCTGCGCAATGATGCTGATGCTGATGCTG 1260  
QY 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 1261 CACTCATTTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 GACTACTGCT 1380  
DB 1321 GACTACTGCT 1380  
QY 1381 GTGCTGCGGCTCCAGCGGGAAGCAACGCGCTGCGGAGTATCCGCGGCAAACTAAGATT 1440  
DB 1381 GTGCTGCGGCTCCAGCGGGAAGCAACGCGCTGCGGAGTATCCGCGGCAAACTAAGATT 1440  
QY 1441 TATGCTCACTGACCAAAACAACAACAACAACAACAACAACAACAACAACAACAACA 1500  
DB 1441 TATGCTCACTGACCAAAACAACAACAACAACAACAACAACAACAACAACAACAACA 1500  
QY 1501 ATATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
DB 1501 ATATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
QY 1561 CTGCTTCAACAGTACCTGCTGACAGCCCTATGAGGAGAGGAGGCTTAAGTCAAGTCA 1620  
DB 1561 CTGCTTCAACAGTACCTGCTGACAGCCCTATGAGGAGAGGAGGCTTAAGTCAAGTCA 1620  
QY 1621 CAACCTGAATGGCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
DB 1621 CAACCTGAATGGCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
QY 1681 CGCCCTCTGCGGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
DB 1681 CGCCCTCTGCGGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779

Db 1741 GTCAGAAATGTCATGCTTGGCTGCGCTACCGATATA 1779

RESULT 2  
AF282887 2096 bp mRNA linear PRI 14-OCT-2000  
LOCUS Homo sapiens heparanase-like protein HPA2c mRNA, complete cds.  
DEFINITION AF282887  
ACCESSION AF282887.1 GI:10801198  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2096)  
McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R.,  
Hirocock, M., Patel, S., Barry, E., Stubberfield, C., Terrett, J. and  
Page, M.  
Cloning and expression profiling of Hpa2, a novel mammalian  
heparanase family member  
Biochem. Biophys. Res. Commun. 276 (3), 1170-1177 (2000)  
2 (bases 1 to 2096)  
McKenzie, E.A., Tyson, K. and Stamps, A.  
Direct Submission  
Submitted (28-JUN-2000) Biology Group, Oxford Glycosciences, 10 The  
Quadrant, Abingdon Science Park, Abingdon OX14 3YS, England  
Location/Qualifiers  
source  
1..2096  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q23-q24"  
61..1839  
/note="splice variant"  
/codon\_start=1  
/product="heparanase-like protein HPA2c"  
/protein\_id="AA023423.1"  
/db\_xref="GI:10801199"  
/translation="MRVLCAPEEAMPSSNRPACALGALYALLLHLSISQAGR  
RPLVDRAGLKEKTLILDVSTKPVTVNENFSLDPSIIHDMFLSKRLV  
TLARGLSPALFPGKRTDFQFONLRPAKSRGPGDYLKRYSDIVASDALDK  
OKGCTIAHPDVMLELOREKAOHMLVLKXOFNSNTSNLITRSJDKYFNADCG  
LHLIPALNLRPNPNSSNSALSILKYSKRYNLSMELGNEPNRYRTMHGAANG  
SGLGDYLOKSLIOPRIYRSASISYGNIGRPKNYVIALDGMKTAGSYDVTMG  
HCYIDGRVYKWDPLKRLDPLSLDQIKRQVNTTTPGKFLHSGVTTISAGTIN  
LSDSYAAGFLWNLTMGLANGIDIVLHSHFDGYNHLVDQNPPLDPLWLSLYR  
LIGPVLAVHVAIGLQRKPRPGRVIRDKLRIYAHCTNMHNHNVAGSITLPIINDHR  
KIKLAGTIRDLKLVHGYLLQYGOEGELKSKSVQANGQPLVWVDGTPELKEPRLRAG  
RLVLPVTMGFFVKNVNALACRR"

ORIGIN

Query Match 99.7%; Score 1774.2; DB 9; Length 2096;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1776; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGGGTGGCTTTGGCTTCCCTGAAGCATGCCCTCAGACAACCTCCGCGCGCG 60  
Db 61 ATGAGGGTGGCTTTGGCTTCCCTGAAGCATGCCCTCAGACAACCTCCGCGCG 120  
QY 61 TGCCTAGCCCGGGGGCTCTCTACTTGGCTGTGTCTCAATCTCTCCCTTCTCCAG 120  
Db 121 TGCCTAGCCCGGGGGCTCTCTACTTGGCTGTGTCTCAATCTCTCCCTTCTCCAG 180  
QY 121 TGCCTAGCCCGGGGGCTCTCTACTTGGCTGTGTCTCAATCTCTCCCTTCTCCAG 180  
Db 121 GGTGAGACAGAGACCCCTGCTGTAGACAGAGTGAAGTTGAAGAAAAGACCTG 180  
QY 121 GGTGAGACAGAGACCCCTGCTGTAGACAGAGTGAAGTTGAAGAAAAGACCTG 180  
Db 181 GGTGAGACAGAGACCCCTGCTGTAGACAGAGTGAAGTTGAAGAAAAGACCTG 240  
QY 181 ATTCTACTTGATGTAGACCAAGAACCCAGTCAAGACATGATGAGAACTTCTCTCT 240

Db 241 ATTCTACTTGATGTAGACCAAGAACCCAGTCAAGACATGATGAGAACTTCTCTCT 300  
QY 241 CTGACGCTGGATCCGTCATCATTTATGATGGCTGATGATTTCTTAAGCTCAAGCGC 300  
Db 301 CTGACGCTGGATCCGTCATCATTTATGATGGCTGATGATTTCTTAAGCTCAAGCGC 360  
QY 301 TTGGTGAACCTTGACCCCGGGGACTTTTGGCCGCTTTCTGCGCTCGGGGGCAAGAGAC 360  
Db 361 TTGGTGAACCTTGACCCCGGGGACTTTTGGCCGCTTTCTGCGCTCGGGGGCAAGAGAC 420  
QY 361 GACTTCCTGAGTTCAGAACCTGAGAACCCCGGGAAGAGCGGGGGCCGGGCGCG 420  
Db 421 GACTTCCTGAGTTCAGAACCTGAGAACCCCGGGAAGAGCGGGGGCCGGGCGCG 480  
QY 421 GATTACTATCTCAAAAAGTATGAGATGACATTTGTCGAAGTATGTTGCTTGAATAA 480  
Db 481 GATTACTATCTCAAAAAGTATGAGATGACATTTGTCGAAGTATGTTGCTTGAATAA 540  
QY 481 CAGAAAGGCTGCAAGATGCGCCGACGACCTGATGTTATGCTGCTCCAAAGGGAAG 540  
Db 541 CAGAAAGGCTGCAAGATGCGCCGACGACCTGATGTTATGCTGCTCCAAAGGGAAG 600  
QY 541 GCAGCTCAGATGATGCTGTTCTTTAAAGAGCAATTCCTCAATCTTACAGTAATCTC 600  
Db 601 GCAGCTCAGATGATGCTGTTCTTTAAAGAGCAATTCCTCAATCTTACAGTAATCTC 660  
QY 601 ATATTTAAGCCAGGCTCTAGACAAACTTTATTAATCTTGTGATGCTCTGGAATCC 660  
Db 661 ATATTTAAGCCAGGCTCTAGACAAACTTTATTAATCTTGTGATGCTCTGGAATCC 720  
QY 661 CTGATATTTGCTTAATGATGATGCTGCTGCTGATTCCTCAATCTTGTGATGCTCTG 720  
Db 721 CTGATATTTGCTTAATGATGATGCTGCTGCTGATTCCTCAATCTTGTGATGCTCTG 780  
QY 721 GCGCTGATGCTGTTGAAGTACAGCGCCGACCAAAAGTAACTTTCTTGGAACTGGGT 780  
Db 781 GCGCTGATGCTGTTGAAGTACAGCGCCGACCAAAAGTAACTTTCTTGGAACTGGGT 840  
QY 781 AATGAGCAAAATACATATGAGATGAGACATGAGCGCGGAGTAAATGGCGGAGTGGGA 840  
Db 841 AATGAGCAAAATACATATGAGATGAGACATGAGCGCGGAGTAAATGGCGGAGTGGGA 900  
QY 841 AAGGATTAATCAATCAAGCTGAAGAGCTGTTGACAGCCATCCGATTTATTCAGAGCGAG 960  
Db 901 AAGGATTAATCAATCAAGCTGAAGAGCTGTTGACAGCCATCCGATTTATTCAGAGCGAG 960  
QY 901 TTATATGGCCCTAATATTTGGGGCGCGAGAGAAAGATGATGCGCCCTCTAGATGATTC 960  
Db 961 TTATATGGCCCTAATATTTGGGGCGCGAGAGAAAGATGATGCGCCCTCTAGATGATTC 1020  
QY 961 ATGAGGTGGCAGAGAGTACAGTATGAGATGAGTTCCTGCAACATTTGCTACATTTGAGC 1020  
Db 1021 ATGAGGTGGCAGAGAGTACAGTATGAGATGAGTTCCTGCAACATTTGCTACATTTGAGC 1080  
QY 1021 CCGGTGCTCAAGTATGAGATTTCTGAAAATCTGCGCTTGTAGACACATCTCTGACCG 1080  
Db 1081 CCGGTGCTCAAGTATGAGATTTCTGAAAATCTGCGCTTGTAGACACATCTCTGACCG 1140  
QY 1081 ATTAGAAAATTCAGAAAATGTTATATCATATCATTCACAGAAAGAAATTTGGCTTGA 1140  
Db 1141 ATTAGAAAATTCAGAAAATGTTATATCATATCATTCACAGAAAGAAATTTGGCTTGA 1200  
QY 1141 GGTGTGATGACCACTCACTGAGGAGCAAAACATCTATCCGATTTCTATGTGACAGA 1200  
Db 1201 GGTGTGATGACCACTCACTGAGGAGCAAAACATCTATCCGATTTCTATGTGACAGA 1260  
QY 1201 TTCTTATGTTGAACATTTAGAAATGCTGGCCAAATCAAGGCAATGATGCTGATACCG 1260  
Db 1261 TTCTTATGTTGAACATTTAGAAATGCTGGCCAAATCAAGGCAATGATGCTGATACCG 1320  
QY 1261 CACTCATTTTGGACATGATGATCAATCACTGCTGAGACCAAGATTTAAACCATTAACA 1320  
Db 1321 CACTCATTTTGGACATGATGATCAATCACTGCTGAGACCAAGATTTAAACCATTAACA 1380





QY 961 ATGAGGTGGCAGAGAGTACAGATGAGTACCTGCGCAACATTGCTACATTGATGAGC 1020  
 DB 1021 ATGAGGTGGCAGAGAGTACAGATGAGTACCTGCGCAACATTGCTACATTGATGAGC 1080  
 QY 1021 CCGGTGGTCAAGGTGATGAGTCTCTGAGAACTGCGCTGTGACACACTCTCTGACGAG 1080  
 DB 1081 CCGGTGGTCAAGGTGATGAGTCTCTGAGAACTGCGCTGTGACACACTCTCTGACGAG 1140  
 QY 1081 ATGAGGAAATTCAGAAAGGTGATATACATACCTCCAGAAAGAGATTGGCTGAA 1140  
 DB 1141 ATGAGGAAATTCAGAAAGGTGATATACATACCTCCAGAAAGAGATTGGCTGAA 1200  
 QY 1141 GGTGGTGGTCAACCTCAGCTGAGGACAAACAACTATCCGATTCCATGCTGACGAG 1200  
 DB 1201 GGTGGTGGTCAACCTCAGCTGAGGACAAACAACTATCCGATTCCATGCTGACGAG 1260  
 QY 1201 TTTCTATGTTGACACTTTAGAAATGCTGGCCAACTAGAGGATGATGCTGATAGG 1260  
 DB 1261 TTTCTATGTTGACACTTTAGAAATGCTGGCCAACTAGAGGATGATGCTGATAGG 1320  
 QY 1261 CACTCAATTTTGTACCATGATGATCAATCACTCTGAGACAGAAATTTTAACTTACCA 1320  
 DB 1321 CACTCAATTTTGTACCATGATGATCAATCACTCTGAGACAGAAATTTTAACTTACCA 1380  
 QY 1321 GACTTACGCT 1380  
 DB 1381 GACTTACGCT 1440  
 QY 1381 GTGGCTGGGCTCCAGCGGAGACCGGCTGGCCGAGTATCCGGACAACTTAAAGATT 1440  
 DB 1441 GTGGCTGGGCTCCAGCGGAGACCGGCTGGCCGAGTATCCGGACAACTTAAAGATT 1500  
 QY 1441 TATGCTCACTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1500  
 DB 1501 TATGCTCACTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1560  
 QY 1501 ATATCAATCTTGATGATGATCAAGAAAGAAATCAAGCTGGGAGCTCTGAGAGCAAG 1560  
 DB 1561 ATATCAATCTTGATGATGATCAAGAAAGAAATCAAGCTGGGAGCTCTGAGAGCAAG 1620  
 QY 1561 CTGGCTTACCAAGTACCTGCTGACGCCCTATGCGGAGAGAGGCTTAAAGTCCAAAGTCA 1620  
 DB 1621 CTGGCTTACCAAGTACCTGCTGACGCCCTATGCGGAGAGAGGCTTAAAGTCCAAAGTCA 1680  
 QY 1621 CAATCAATGCGCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB 1681 CAATCAATGCGCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 QY 1681 CCGCCCTTGGGCGCGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
 DB 1741 CCGCCCTTGGGCGCGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 QY 1741 GTCAAGAAATGTCATGCTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1779  
 DB 1801 GTCAAGAAATGTCATGCTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1839

RESULT 4  
 AX180830 2636 bp DNA linear PAT 06-AUG-2001  
 DEFINITION Sequence 1 from Patent WO0146392.  
 ACCESSION AX180830  
 VERSION AX180830.1 GI:15132666  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 McKenzle, E.A., Stamps, A.C., Terrell, J.A. and Tyson, K.L.  
 Homologies of human heparanase and splice variants thereof  
 Patent: WO 0146392-A 1 28-JUN-2001.

## Oxford Glycosciences (UK) Limited (GB)

## FEATURES

Location/Qualifiers  
 1..2636  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 601..2379  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC50008.1"  
 /db\_xref="GI:15132667"  
 /db\_xref="EMBL:CAC50008"  
 /translation="MRVLCAPFEMPSNSRSPACAPALYALALHLSLSSQAGDR  
 RPLPVDPAAGLKEKTLILDLSTKPNPVATNENFLSLQDPSIIHGMWDLPSKRLV  
 TLARGLSPAFLEFGSKTDLOFONLBNPAKSRGPDYLLKNVEDIVRSVALDK  
 OKGCKLAHQDVMLEORBAQOMHVLLEKPSNTYSNLIILARSIDLKYNFADCG  
 LHIIFNLNLRBNPNWSXSSALSLIKYASAKKXNLSWELGNEPNVPTMRVAVNG  
 SOLGKDYIKSLIOPRIATSPASISVGNIPGRPRNVALLDSPFMYASVDAVTVQ  
 HCTIDRIVKVMDFLTKRLDITLSDQIRKIQRVNTTPEKTLMBGVTTSGTIN  
 LHSYVAGFLMLTLGLANQIDVIRHSPFHGYNHLVDQNFNPLDPLWLSLYR  
 LIQPKVLAVAVAGLQKRPGRVIRDKLRIVACTNHNHNRYVGSITFLINLHRSR  
 KKLKAGTLDKLVHQLDLPYQSGKSKSVQNLNGPLVWVDGILPELKPPLPAG  
 RLVIPFVITGPFVKNVNALCRYR"

## CDS

## ORIGIN

Query Match 99.7%; Score 1773.8; DB 6; Length 2636;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1775; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGGGTGGCTTTGTGCTTCCCTGAAGCCATGCGCTCCAGCACTCCCGCCCCCGCG 60  
 DB 601 ATGAGGGTGGCTTTGTGCTTCCCTGAAGCCATGCGCTCCAGCACTCCCGCCCCCGCG 660  
 QY 61 TGCTTACGCGCGGGGGCTCTCTACTTGGCTCTGTGCTGTCATCTCTCCCTTCTCCAG 120  
 DB 661 TGCTTACGCGCGGGGGCTCTCTACTTGGCTCTGTGCTGTCATCTCTCCCTTCTCCAG 120  
 QY 121 GCTGAGACAGAGAGACCTTGCGCTGTGAGACAGCTGACAGTTTGAAGAAAGACCTG 180  
 DB 721 GCTGAGACAGAGAGACCTTGCGCTGTGAGACAGCTGACAGTTTGAAGAAAGACCTG 180  
 QY 721 GCTGAGACAGAGAGACCTTGCGCTGTGAGACAGCTGACAGTTTGAAGAAAGACCTG 780  
 DB 781 ATTTACTTATGATGAGACCAAGAACCAAGTCAAGACAGTCAATGAGAACTTCTCTCT 240  
 QY 241 CTGAGAGTGAATCGTCCATCATCATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 841 CTGAGAGTGAATCGTCCATCATCATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 841 CTGAGAGTGAATCGTCCATCATCATGATGATGATGATGATGATGATGATGATGATGAT 900  
 DB 901 TTGGTGAACCTTGCGCGGGGACTTTCGCGCCCTTTCGCGCTTGGCGGGCAAAAGAGAC 360  
 QY 361 GACTTCTGCAAGTTCAGAACTGAGAGAACCGCGCGGAGAGCGCGCGGGCGCGCGCG 420  
 DB 961 GACTTCTGCAAGTTCAGAACTGAGAGAACCGCGCGGAGAGCGCGCGGGCGCGCGCG 1020  
 QY 421 GATTACTATCTCAAAAACATATGAGATGATGATGATGATGATGATGATGATGATGATG 480  
 DB 1021 GATTACTATCTCAAAAACATATGAGATGATGATGATGATGATGATGATGATGATGATG 1080  
 QY 481 CAGAAAGGCTGCAAGATTTGCCGACAGCCTGATGATGATGATGATGATGATGATGATG 540  
 DB 1081 CAGAAAGGCTGCAAGATTTGCCGACAGCCTGATGATGATGATGATGATGATGATGAT 1140  
 QY 541 GCAAGTCAAGTCAATGCTTCTTCTTAAAGAGCAATTTCCAACTTACAGTAATCTC 600  
 DB 1141 GCAAGTCAAGTCAATGCTTCTTCTTAAAGAGCAATTTCCAACTTACAGTAATCTC 1200  
 QY 601 ATATTAAAGCCAGAGTCTTGAACAACTTTATATACCTTGTGCTGATGCTGCACTCCAC 660  
 DB 1201 ATATTAAAGCCAGAGTCTTGAACAACTTTATATACCTTGTGCTGATGCTGCACTCCAC 1260

QY 661 CTGATATTGCTCTAAATGACATGCGTCGTAATCCCAATTAATCTCTGGAGACGTTCTAGT 720  
 Db 1261 CTGATATTGCTCTAAATGACATGCGTCGTAATCCCAATTAATCTCTGGAGACGTTCTAGT 1320  
 QY 721 GCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAAGTACACATTTCTTGGAACTGGGT 780  
 Db 1321 GCCCTGAGTCTGTTGAAGTACAGCGCCAGCAAAAAGTACACATTTCTTGGAACTGGGT 1380  
 QY 781 AATGAGCAATTAATGATGAGACCATGATGCGCGGAGTAAATGAGCAGCAGTGGGA 840  
 Db 1381 AATGAGCAATTAATGATGAGACCATGATGCGCGGAGTAAATGAGCAGCAGTGGGA 1440  
 QY 841 AAGGATTACATCCAGCTGAAGAGCTGTTGACGCCCATCCGATTTATTTCCAGAGCCAGC 900  
 Db 1441 AAGGATTACATCCAGCTGAAGAGCTGTTGACGCCCATCCGATTTATTTCCAGAGCCAGC 1500  
 QY 901 TATATGCGCTTAATTTGGGCGCGGAGAGAAATGTCATGCGCTCTCTAGATGATTC 960  
 Db 1501 TATATGCGCTTAATTTGGGCGCGGAGAGAAATGTCATGCGCTCTCTAGATGATTC 1560  
 QY 961 ATGAAAGTGGCAGGAAGTACAGTATGAGTACCTGCGCAACATGCTACATTTAGTGC 1020  
 Db 1561 ATGAAAGTGGCAGGAAGTACAGTATGAGTACCTGCGCAACATGCTACATTTAGTGC 1620  
 QY 1021 CCGGTGCTAGAGTGTATGACCTTCCGAAAACCTCGCTGTTAGACACCTCTGACAG 1080  
 Db 1621 CCGGTGCTAGAGTGTATGACCTTCCGAAAACCTCGCTGTTAGACACCTCTGACAG 1680  
 QY 1081 ATTAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGATTTGGCTTGA 1140  
 Db 1681 ATTAGAAATTCAGAAAGTGTATTAATACATCACTCCAGAAAGATTTGGCTTGA 1740  
 QY 1141 GGTGTGTGACCACTGACCTGAGGAGCAAAAACATCTATCCGATTTCTTACTGACAGA 1200  
 Db 1741 GGTGTGTGACCACTGACCTGAGGAGCAAAAACATCTATCCGATTTCTTACTGACAGA 1800  
 QY 1201 TTCTTATGTTGAACCTTTAAGAAATGCTGGCAATCAAGGAGTATGTCGTATGAG 1260  
 Db 1801 TTCTTATGTTGAACCTTTAAGAAATGCTGGCAATCAAGGAGTATGTCGTATGAG 1860  
 QY 1261 CACTATTTTGTGACCATGATGATCAATCACTCTGTGTGACCAAGATTTTAACTTACCA 1320  
 Db 1861 CACTATTTTGTGACCATGATGATCAATCACTCTGTGTGACCAAGATTTTAACTTACCA 1920  
 QY 1321 GACTATGAGTCTCTCTCTCTACAAAGCCCTGATGAGCCCAAGTCTTGGCTGTGAT 1380  
 Db 1921 GACTATGAGTCTCTCTCTCTCTACAAAGCCCTGATGAGCCCAAGTCTTGGCTGTGAT 1980  
 QY 1381 GTGGCTGGGCTCCAGGGAAGCCAGGCTGCGCAAGTATCCGAGCAAACTAAGAT 1440  
 Db 1981 GTGGCTGGGCTCCAGGGAAGCCAGGCTGCGCAAGTATCCGAGCAAACTAAGAT 2040  
 QY 1441 TATGCTCATGCAAAACCAACACCAACCACTACGTTCTGTGGTCAATTCATCTTTT 1500  
 Db 2041 TATGCTCATGCAAAACCAACACCAACCACTACGTTCTGTGGTCAATTCATCTTTT 2100  
 QY 1501 ATCATCACTGTCATGATCAAGAAAGAAATCAAGCTGTGGAGCTCTCAGAGCAAG 1560  
 Db 2101 ATCATCACTGTCATGATCAAGAAAGAAATCAAGCTGTGGAGCTCTCAGAGCAAG 2160  
 QY 1561 CTGGTTCAACAGTACCTGCTGAGGCTTATGGGAGAGAGGCTTAAAGTCCAAAGCAAG 1620  
 Db 2161 CTGGTTCAACAGTACCTGCTGAGGCTTATGGGAGAGAGGCTTAAAGTCCAAAGCAAG 2220  
 QY 1621 CAACTGAATGCGCAGCCCTTAAGTATGATGAGAGAGAGCCCTCCAGAAATGGAAGCC 1680  
 Db 2221 CAACTGAATGCGCAGCCCTTAAGTATGATGAGAGAGAGCCCTCCAGAAATGGAAGCC 2280  
 QY 1681 CCGCCCTTCCGCGCGGAGCATGTCATCCCTCCAGTACACATGCGCTTTATATG 1740  
 Db 2281 CCGCCCTTCCGCGCGGAGCATGTCATCCCTCCAGTACACATGCGCTTTATATG 2340  
 QY 1741 GTCAAGATGTCATGCTTGGCTGCGCTACCGATTA 1779

Db 2241 GTCAAGATGTCATGCTTGGCTGCGCTACCGATTA 2379

RESULT 5  
 AX393920  
 LOCUS  
 DEFINITION  
 AX393920  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AX393920 1847 bp DNA linear PAT 23-MAR-2002  
 Sequence 1 from Patent WO204645.  
 AX393920  
 AX393920.1 GI:19701880

REFERENCE  
 1  
 David, G. and Duerr, J.  
 A second human heparanase, and splice variants thereof, with a  
 predominant expression in skeletal muscle, heart and pancreas  
 Patent: WO 0204645-A 1 17-JAN-2002;  
 Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BE)

FEATURES  
 source  
 1. 1847  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 85.1833  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD29059.1"  
 /db\_xref="GI:19701881"  
 /db\_xref="REFSEQ:CAD29059"  
 /translation="MPSNSRPRACAPALVIALIHLHSLSSQADRRPLPVPDAG  
 LKRTLLIDVSTNRPVTVNRFSLQDPSITHGMDIFLSKRLVTLARGLSPF  
 LRFGGRADFLQNLNRPAPKSGGPGPYLYLKNEDDIVRDVADLQKQKGLIAP  
 DWMLREKRLAQMLVLLKEQSNSTNSMLITARLDLXVNPADCSGLIFALNAL  
 RPNPNNSNNSLSLKLKYSASAKKNI SWELQGMKPNVYRMTGKRAVNSGLIDGYOL  
 KSLLOPRLIYRSASLYGPNIGRKRKIVLALDGMKAVAGTVDATVMQHCYIDGRVVK  
 VMDLFTLLDITSDQIRIKIQRVNTYEGKKIMEGVTTSGAGTNNLSDSYAAGEL  
 WNTLFLMLANOGIDIVIRSRSPHGYNHVNDQNPFLPDVMTSLYKRLIGPVLAHV  
 VAGLQKRPGRVIRDKRIYIAHCTNHNHNVRSGLITFINLHRSKKIKLAGTLR  
 DKVHOYLLOPYGSGSKSVQNLQNPVMDGTLPELKPRLPLAAGRTLVIPVMT  
 GPEVVAHVNALACRYR"

ORIGIN  
 Query Match 99.4%; Score 1769; DB 6; Length 1847;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1772; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

1 ATGAGGGTGTCTTGTGCTTCCCTGAAGCCATCCCTCCAGCAACTCCGCGCCCGCG 60  
 55 ATGAGGGTGTCTTGTGCTTCCCTGAAGCCATCCCTCCAGCAACTCCGCGCCCGCG 114  
 61 TGGCTAGCCCGGGGGGCTCTACTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 115 TGGCTAGCCCGGGGGGCTCTACTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 174  
 121 GCTGAGAGCAGAGACCCCTTGCCTGAGACAGAGCTGCAAGTTTGAAGAAAAGACCTTG 180  
 175 GCTGAGAGCAGAGACCCCTTGCCTGAGACAGAGCTGCAAGTTTGAAGAAAAGACCTTG 234  
 181 ATTCTACTGATGAGACCAAGAACCCAGTACAGAGCAAGTCAATAGAACTTCTCTCT 240  
 235 ATTCTACTGATGAGACCAAGAACCCAGTACAGAGCAAGTCAATAGAACTTCTCTCT 294  
 241 CTGAGAGTGAATCCGTCATCATTCATGATGATGCTGCTGATTTCTTAAGCTTCAAGCGC 300  
 295 CTGAGAGTGAATCCGTCATCATTCATGATGATGCTGCTGATTTCTTAAGCTTCAAGCGC 354  
 301 TTGGTGAATCCGTCATCATTCATGATGATGCTGCTGATTTCTTAAGCTTCAAGCGC 360  
 355 TTGGTGAATCCGTCATCATTCATGATGATGCTGCTGATTTCTTAAGCTTCAAGCGC 414

QY	36	GACTTCCTGAGATTCCGAAACCTGAGGAACCCGGCGAAAAAGCCGCGGAGCCCGCGGACCCG	420
Db	415	GACTTCCTGAGATTCCGAAACCTGAGGAACCCGGCGAAAAAGCCGCGGAGCCCGCGGACCCG	474
QY	421	GATTACTATCTCAAAAACTATGAGATGACATTGTCGAAAGTAATGTTCTCTTAGATPAA	480
Db	475	GATTACTATCTCAAAAACTATGAGATGACATTGTCGAAAGTAATGTTCTCTTAGATPAA	534
QY	481	CAGAAAGGCTGCAGAAATTGGCCCGACCCCTGAGTATATGCTGGTGCTCCAAAGGAGAA	540
Db	535	CAGAAAGGCTGCAGAAATTGGCCCGACCCCTGAGTATATGCTGGTGCTCCAAAGGAGAA	594
QY	541	GCAGCTCAGATGCATCTGTTCTCTTAAAGAGCAATTCACAACTTACAGTAATCTC	600
Db	595	GCAGCTCAGATGCATCTGTTCTCTTAAAGAGCAATTCACAACTTACAGTAATCTC	654
QY	601	ATATTTAAACGCGAGGCTCTTAGACAACTTTATATCTTTGCTGATGCTCTGACATCCAC	660
Db	655	ATATTTAAACGCGAGGCTCTTAGACAACTTTATATCTTTGCTGATGCTCTGACATCCAC	714
QY	661	CTGATTTTGGCTAAATGACCTGCGATGATCCCATTAATCCTGGAACAGTCTAGT	720
Db	715	CTGATTTTGGCTAAATGACCTGCGATGATCCCATTAATCCTGGAACAGTCTAGT	774
QY	721	GCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAAGTACAACTTTCTTGGAACTGGGT	780
Db	775	GCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAAGTACAACTTTCTTGGAACTGGGT	834
QY	781	AATAGGCAAAATATCTATCCGACCATGATGAGCGCGGCAATPAAATGGCAGCAGTTGGGA	840
Db	835	AATAGGCAAAATATCTATCCGACCATGATGAGCGCGGCAATPAAATGGCAGCAGTTGGGA	894
QY	841	AAGATTATCATCCAGCTGGAAGAGCCTGTGGAGCCATCCGGATTATTCGAGCCAGC	900
Db	895	AAGATTATCATCCAGCTGGAAGAGCCTGTGGAGCCATCCGGATTATTCGAGCCAGC	954
QY	901	TTTATATGGCCCTAATTTTGGCGCGCCAGAGAAATGATATCGCCCTCTAGATGGAATC	960
Db	955	TTTATATGGCCCTAATTTTGGCGCGCCAGAGAAATGATATCGCCCTCTAGATGGAATC	1014
QY	961	ATGAAGGTGAGAGATPACAGTATGATGCAATTACTCTGGCAACATTCGTACATTGATGC	1020
Db	1015	ATGAAGGTGAGAGATPACAGTATGATGCAATTACTCTGGCAACATTCGTACATTGATGC	1074
QY	1021	CGGCTGTCAAGGTGATGAGCTTCTCTGAAAACTGCGCTGTTAGACAACACTCTTGACACAG	1080
Db	1075	CGGCTGTCAAGGTGATGAGCTTCTCTGAAAACTGCGCTGTTAGACAACACTCTTGACACAG	1134
QY	1081	ATTAGGAATAATTCAGAAAGGTGTTTATATCATATCACTCCAGGAAAGAAAGATTGGCTTGA	1140
Db	1135	ATTAGGAATAATTCAGAAAGGTGTTTATATCATATCACTCCAGGAAAGAAAGATTGGCTTGA	1194
QY	1141	GGTGTGTGACCACTCGAGCTGGAGGCAAAACAATCTATTCGATTCTTATGCTGCAGGA	1200
Db	1195	GGTGTGTGACCACTCGAGCTGGAGGCAAAACAATCTATTCGATTCTTATGCTGCAGGA	1254
QY	1201	TTTCTATAGTTGAAACATTTGGAGATGCTGGCCATCAAGGGCAATTGATGTGTATTCGG	1260
Db	1255	TTTCTATAGTTGAAACATTTGGAGATGCTGGCCATCAAGGGCAATTGATGTGTATTCGG	1314
QY	1261	CACCTCATTTTTGACCATGATATCAATCACTCGTGAGCAAGAAATTTTAACTTACCATTACA	1320
Db	1315	CACCTCATTTTTGACCATGATATCAATCACTCGTGAGCAAGAAATTTTAACTTACCATTACA	1374
QY	1321	GACTACTGGCTCTCTCTCTCTTACAAAGGCGCTGATCCGCGCCCAAGTCTTGGCTGTGAT	1380
Db	1375	GACTACTGGCTCTCTCTCTCTTACAAAGGCGCTGATCCGCGCCCAAGTCTTGGCTGTGAT	1434
QY	1381	GTGGCTGGGCTCAACGGAAGCAAGGCTGGCCAGATGATCCGGAACAACTAGAGATT	1440
Db	1435	GTGGCTGGGCTCAACGGAAGCAAGGCTGGCCAGATGATCCGGAACAACTAGAGATT	1494
QY	1441	TATGCTCACTGCACAAACCAACAACAACAACATCTAGTTCTGGGGTCCATTACACTTTT	1500

Db	1495	TATGTCACGTGCACAACCAACCACAACCAACTAAGTTCGTGGTGCCATTACACTTTT	1554
Qy	1501	ATATCAACTGTGATGATCAAGAAGAAATACTAACCTGTGGGACTCTCAGAGACAAG	1560
Db	1555	ATATCAACTGTGATGATCAAGAAGAAATACTAACCTGTGGGACTCTCAGAGACAAG	1614
Qy	1561	CTGTTTACCACGATGACCTGTGTCGACCCCTATGGCGACAGAGGCCCTTAAGTCCAGTCAGTG	1620
Db	1615	CTGTTTACCACGATGACCTGTGTCGACCCCTATGGCGACAGAGGCCCTTAAGTCCAGTCAGTG	1674
Qy	1621	CAACTGAATGGCCAGCCCTTAGTGATGGTGGAGAGAGAGGACCTCCCAGAAATTGAAGCCC	1680
Db	1675	CAACTGAATGGCCAGCCCTTAGTGATGGTGGAGAGAGAGGACCTCCCAGAAATTGAAGCCC	1734
Qy	1681	CGCCCCCTTGGGGCCGGCCGGAATTCATGTCATCCCTCCAGTCCACATGGGCTTTTAGTG	1740
Db	1735	CGCCCCCTTGGGGCCGGCCGGAATTCATGTCATCCCTCCAGTCCACATGGGCTTTTAGTG	1794
Qy	1741	GTCAGAAATGTCATATGCTTTGGSCCTTGCCGCTACCGATPA	1779
Db	1795	GTCAGAAATGTCATATGCTTTGGSCCTTGCCGCTACCGATPA	1833
RESULT 6			
HSA299720			
LOCUS	HSA299720	1872 bp	mRNA linear PRI 04-JAN-2002
DEFINITION	Homo sapiens mRNA for heparanase 3.		
ACCESSION	AJ299720		
VERSION	AJ299720.1 GI:18073438		
KEYWORDS	heparanase 3.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Legoux,P., Legoux,R., O'Brien,D. and Salome,M.		
JOURNLT	Unpublished		
REFERENCE	2 (bases 1 to 1872)		
AUTHORS	Pessegue Safontas,B.J.O.P.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-SEP-2000) Pessegue Safontas B.J.O.P.S., Bioinformatics, Sanoofi-SyntheLabo, Sanoofi Recherche - innopole - Vote 1 - BP 137, 31676 Labège CEDEX, FRANCE		
FEATURES	Location/Qualifiers		
Source	1..1872		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue type="prostate"		
CDS	1..1647		
	/codon_start=1		
	/product="heparanase 3"		
	/protein_id="OAc82452.1"		
	/db_xref="GI:18073439"		
	/db_xref="SPTRMBL:O8MWOL"		
	/translation="MRVLCAPPEAMLSNSRPECIAPGLAYIALTLISLSQAGDR RPUPVDRAAGKEKTEILLIDPSTENIPRYVENFSLSDLPSSIHDEMDPLFSKRIV TIARGSAPFLRFGRGKRTDFLOFONLHPAKSGSGPGPDYILKNVEDDIVRSVALDK QKRGASQHPDWLTELORERKAOMHLVLKEOPSNTYSNLITLASLDKLXNSADCSG LHLIFALNALRRPNPNRNSSSALSILTKYASASKYNI SWELGNPPNNRYTMHBRAYWG SQGKDVIQKSLDPLFTIRLRSLASVIGPNIGRPKNVIALLDGEPKAGSVDAVTWQ HCYIKDVVRVKMDLFKTRTLRLDTLSDOIIRKI QKVNVYTPBCKTIMLBGVVTSAGTNV LBSVSYAAGFLMTLMANQAGVIDVIRHSFPDHGYNHAVDQNFNPDPWLSLYKR LIGKYLVAIVAGIQREKPRGRVIRDGLRIYATCNHNHNHYVSGITLFI INLHRSR KKIKLAGTLRDKLIVHQYLQPYGEGSKSYKGCOYCII"		
ORIGIN			
Query Match	95.2%	Score 1694.2;	DB 9; Length 1872;
Best Local Similarity	96.1%;	Pred. No. 0;	
Matches 1776;	Conservative 0;	Mismatches 3;	Indels 70; Gaps 1;
1 ATGAGGGGTCCTTGTGCTTCCCTGAGCATGCGCTTCAGACAATCCTCCGCCCCCGCG 60			

Db	1	ATGAGGCTGCTTTGCTTCCCTGAGGCACTCTCCAGCACTCCGCCCCCGG	60
Qy	61	TGCTAGCCCCGGGGGCTCTCTAATTGGCTTGTGCTCACTTCCCTTTCTCCAG	120
Db	61	TGCTAGCCCCGGGGGCTCTCTAATTGGCTTGTGCTCACTTCCCTTTCTCCAG	120
Qy	121	GCCTGAGACAGAGACCTTGCCTGTAGACAGAGCTGACAGTTTGAAGAAAAGCCCTG	180
Db	121	GCCTGAGACAGAGACCTTGCCTGTAGACAGAGCTGACAGTTTGAAGAAAAGCCCTG	180
Qy	181	ATTCTACTGTATGTAGACCAAGAACCCAGTCAAGACAGTCAATGAGAATTCTCTCT	240
Db	181	ATTCTACTGTATGTAGACCAAGAACCCAGTCAAGACAGTCAATGAGAATTCTCTCT	240
Qy	241	CTGACAGTGAATCCGTTCATCATCATGATGAGCTGAGTGTGCTTCCCTAGCTCCAG	300
Db	241	CTGACAGTGAATCCGTTCATCATCATGATGAGCTGAGTGTGCTTCCCTAGCTCCAG	300
Qy	301	TTGGTGAACCTGAGCCCGGGGACTTTCCGCTTTCTGCGCTTCCGCGGCAAAAGGAC	360
Db	301	TTGGTGAACCTGAGCCCGGGGACTTTCCGCTTTCTGCGCTTCCGCGGCAAAAGGAC	360
Qy	361	GACTTCTGCACTTCCAGAACTGAGAGACCCGCGGAAAGCCGCGGGGCGCGGCGG	420
Db	361	GACTTCTGCACTTCCAGAACTGAGAGACCCGCGGAAAGCCGCGGGGCGCGGCGG	420
Qy	421	GATTACTATCTCAAAAATATGAGATGACATTTGTGAGATGATGTTGCTTGAATAA	480
Db	421	GATTACTATCTCAAAAATATGAGATGACATTTGTGAGATGATGTTGCTTGAATAA	480
Qy	481	CAGAAAGGCTGCAAGATTTGCCAGACACCTGATGTATGCTGAGCTCCAAAGGAGAA	540
Db	481	CAGAAAGGCTGCAAGATTTGCCAGACACCTGATGTATGCTGAGCTCCAAAGGAGAA	540
Qy	541	GCAGTCAAGTCACTGCTTCTCTTAAGAGCAATTCATCTACATCTCACTCACTC	600
Db	541	GCAGTCAAGTCACTGCTTCTCTTAAGAGCAATTCATCTACATCTCACTCACTC	600
Qy	601	ATATTACAGCAAGCTCTCTAGACAACTTATTAATCTGCTGATCTCTGAGCTCCAC	660
Db	601	ATATTACAGCAAGCTCTCTAGACAACTTATTAATCTGCTGATCTCTGAGCTCCAC	660
Qy	661	CTGATATTTGCTTGAAGTACAGCGCCGCAAAAAGTACAACTTTCTTGGGAACTGGGT	720
Db	661	CTGATATTTGCTTGAAGTACAGCGCCGCAAAAAGTACAACTTTCTTGGGAACTGGGT	720
Qy	721	GCCCTGAGTCTGTGAGTACAGCGCCGCAAAAAGTACAACTTTCTTGGGAACTGGGT	780
Db	721	GCCCTGAGTCTGTGAGTACAGCGCCGCAAAAAGTACAACTTTCTTGGGAACTGGGT	780
Qy	781	AATGAGCAATTAATATGAGCACTGATGCGCGGAGTAAATGAGAGCACTTGGGA	840
Db	781	AATGAGCAATTAATATGAGCACTGATGCGCGGAGTAAATGAGAGCACTTGGGA	840
Qy	841	AAGGATTAATCACTCACTGAGAGGCTGTGAGAGCCCATCCGATTTATTCAGAGCCAG	900
Db	841	AAGGATTAATCACTCACTGAGAGGCTGTGAGAGCCCATCCGATTTATTCAGAGCCAG	900
Qy	901	TTATATGCGCTTAATATGAGGCGCGAGAGAAATGATGCGCCCTCTAGATGAGATTC	960
Db	901	TTATATGCGCTTAATATGAGGCGCGAGAGAAATGATGCGCCCTCTAGATGAGATTC	960
Qy	961	ATGAGTGTGAGAGAGTACAGTATGAGTCACTTGGGCAACTTGTCACTTGAATGGC	1020
Db	961	ATGAGTGTGAGAGAGTACAGTATGAGTCACTTGGGCAACTTGTCACTTGAATGGC	1020
Qy	1021	CGGCTGTGAGAGTATGAGTCTTCTGAAATCGCCCTGTATGAGACACTCTTGAACAG	1080
Db	1021	CGGCTGTGAGAGTATGAGTCTTCTGAAATCGCCCTGTATGAGACACTCTTGAACAG	1080
Qy	1081	ATTAGAAAATTCAGAAAGTGTATATCATCACTCCAGAGAAAGATTTGGCTTGA	1140
Db	1081	ATTAGAAAATTCAGAAAGTGTATATCATCACTCCAGAGAAAGATTTGGCTTGA	1140
Qy	1141	GGTGTGAGACCTCACTGAGTGAAGGCAAAAGATCATGCTTCTATGCTGAGGA	1200
Db	1141	GGTGTGAGACCTCACTGAGTGAAGGCAAAAGATCATGCTTCTATGCTGAGGA	1200
Qy	1201	TTCTTATGTTGAACACTTTAGATGCTGCAATGAGGCAATGATGCTGATACGG	1260
Db	1201	TTCTTATGTTGAACACTTTAGATGCTGCAATGAGGCAATGATGCTGATACGG	1260
Qy	1261	CAGTATTTTGTGACATGATGATCAATCACTCGTGAACCGAATTTTAAACCAATCA	1320
Db	1261	CAGTATTTTGTGACATGATGATCAATCACTCGTGAACCGAATTTTAAACCAATCA	1320
Qy	1321	GACTATGCTCTCTCTCTCTCAAGAGCGCTGATGCGCCCAAAATCTTGGCTGTGAT	1380
Db	1321	GACTATGCTCTCTCTCTCTCTCAAGAGCGCTGATGCGCCCAAAATCTTGGCTGTGAT	1380
Qy	1381	GTTGCTGAGCTTCCAGCGGAAAGCCAGCGCTGCGGAGTATCCGCGAATAAGATT	1440
Db	1381	GTTGCTGAGCTTCCAGCGGAAAGCCAGCGCTGCGGAGTATCCGCGAATAAGATT	1440
Qy	1441	TATGCTCACTGACAAACCAACCAACCAACTAGTTCGTGGGTCCATTACTTTT	1500
Db	1441	TATGCTCACTGACAAACCAACCAACCAACTAGTTCGTGGGTCCATTACTTTT	1500
Qy	1501	ATCATCACTTGCATCATCAAGAAAGAAATCAAGCTGCTGGAGCTCTGAGAGCAAG	1560
Db	1501	ATCATCACTTGCATCATCAAGAAAGAAATCAAGCTGCTGGAGCTCTGAGAGCAAG	1560
Qy	1561	CTGTTTCACTGATCTCTCTGAGCTGAGCTTATGAGGAGAGGCTTAACTCAAAACCA	1620
Db	1561	CTGTTTCACTGATCTCTCTGAGCTGAGCTTATGAGGAGAGGCTTAACTCAAAACCA	1620
Qy	1614	-----	1613
Db	1621	AGATGTCATATCTGAGATCATCTGAATCTTAAAGGCACTTAACTTCTTCAAGAA	1680
Qy	1614	-----GTCAAGTCAACTGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1670
Db	1681	AGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1740
Qy	1671	ATTGAGCCCGCCCGCTTGGGCGGCGGAGCACTTGTATCTCTCAAGTCACTGAG	1730
Db	1741	ATTGAGCCCGCCCGCTTGGGCGGCGGAGCACTTGTATCTCTCAAGTCACTGAG	1800
Qy	1731	CTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1779
Db	1801	CTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1849
RESULT 7			
AX393924			
LOCUS 1685 bp DNA linear PAT 23-MAR-2002			
DEFINITION Sequence 5 from Patent WO0204645.			
ACCESSION AX393924			
VERSION AX393924.1 GI:19701884			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE 1 David, G. and Duerr, J.			
AUTHORS A second human heparanase, and splice variants thereof, with a			
TITLE predominant expression in skeletal muscle, heart and pancreas			
JOURNAL Patent: WO 0204645-A 5 17-JAN-2002;			
FEATURES Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BE)			
location/Qualifiers			
1..1685			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			

CDS

85..1671  
 /note="unnamed protein product"  
 /protein\_id="CAD29061.1"  
 /db\_xref="gi:19701885"  
 /db\_xref="remtbl:CAD29061"  
 /translation="MPSNSRPACTAGALYLAALLHLSSOAGDRRLPVDRAG  
 LKEXTLIDVSTKNPRTVNENPLSDPSIIHDEWLDPLSSKRLVTLARGLSPAF  
 LAFGKRRADFLQFONLRNPAKSRGSGPDYLYKVEPARSLDKYNPADCSGLTIRA  
 LNALRNPNSWNSALSLLKTSASKRTNLSWELGPNPNRTMGRANVSGLDG  
 YVLSLQPIRISYRSLYGNIGRPRKRVIALDGMKVASITDAVTVQHTYIG  
 RYVKTMDPKRLDLDLSIDQIRIKQVNTYTPGKXMLEVSVTSAGTNLSDSA  
 AGFLMNTGLMANGIDIVLIRHSFDEGHNHLDONNPLPDYWSILYRLLGPVY  
 LAVHAGLQKRPGRVIRADKLIRYACGNNHNNHYVSGSLTFLINLRSRKKIKLA  
 GTLRKLVHQLVLPYGGGGLKSKSVQNLQDPLVWVDGTLPELKPRLRAGRTLVLP  
 PVTMGFPVRYKVNALACRRR"

ORIGIN

Query Match 81.0%; Score 1441.4; DB 6; Length 1685;  
 Best Local Similarity 90.7%; Pred. No. 0;  
 Matches 1614; Conservative 1; Mismatches 2; Indels 162; Gaps 1;

QY 1 ATGAGGGTCTTTTTCCTTCCCTGAAGCCATGCCCTCAGAACTCCGCGCCCGCGCG 60  
 DB ATGAGGGTCTTTTTCCTTCCCTGAAGCCATGCCCTCAGAACTCCGCGCCCGCG 114  
 QY 61 TGGCTAGCCCGGGGCTCTACTTGGCTGTGCTCATCTCTCCCTTCTCCGAG 120  
 DB TGGCTAGCCCGGGGCTCTACTTGGCTGTGCTCATCTCTCCCTTCTCCGAG 174  
 QY 115 TGGCTAGCCCGGGGCTCTACTTGGCTGTGCTCATCTCTCCCTTCTCCGAG 174  
 DB TGGCTAGCCCGGGGCTCTACTTGGCTGTGCTCATCTCTCCCTTCTCCGAG 234  
 QY 121 GCTGAGACAGAGACCTTGGCTGTGACAGAGCTGCAAGTTTGAAGAAAAGCCTG 180  
 DB GCTGAGACAGAGACCTTGGCTGTGACAGAGCTGCAAGTTTGAAGAAAAGCCTG 234  
 QY 175 GCTGAGACAGAGACCTTGGCTGTGACAGAGCTGCAAGTTTGAAGAAAAGCCTG 234  
 DB GCTGAGACAGAGACCTTGGCTGTGACAGAGCTGCAAGTTTGAAGAAAAGCCTG 240  
 QY 181 ATCTCTAGTGTGACACCAAGACCAAGTCAAGAGTCAATGAGAACTCTCT 240  
 DB ATCTCTAGTGTGACACCAAGACCAAGTCAAGAGTCAATGAGAACTCTCT 294  
 QY 235 ATCTCTAGTGTGACACCAAGACCAAGTCAAGAGTCAATGAGAACTCTCT 294  
 DB ATCTCTAGTGTGACACCAAGACCAAGTCAAGAGTCAATGAGAACTCTCT 354  
 QY 241 CTGACAGTGTATCCGTCATCATTCATGATGGCTGCTGATTTCTTAAGCTCAAGCG 300  
 DB CTGACAGTGTATCCGTCATCATTCATGATGGCTGCTGATTTCTTAAGCTCAAGCG 354  
 QY 301 TTGGTAACTGGCCCGGGGACTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGGAGC 360  
 DB TTGGTAACTGGCCCGGGGACTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGGAGC 414  
 QY 355 TTGGTAACTGGCCCGGGGACTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGGAGC 414  
 DB TTGGTAACTGGCCCGGGGACTTTGGCCGCTTCTGCGCTTGGGGGCAAAAGGAGC 474  
 QY 361 GACTCTGCACTTCAGAACTGAGAAACCGGGCGAAAGCGCGGGGCGCGGCGCG 420  
 DB GACTCTGCACTTCAGAACTGAGAAACCGGGCGAAAGCGCGGGGCGCGGCGCG 474  
 QY 415 GACTCTGCACTTCAGAACTGAGAAACCGGGCGAAAGCGCGGGGCGCGGCGCG 474  
 DB GACTCTGCACTTCAGAACTGAGAAACCGGGCGAAAGCGCGGGGCGCGGCGCG 480  
 QY 421 GATTACTATCTCAAAACCTATGAGATGATTTGTTGAAAGTGTGCTTAATAA 480  
 DB GATTACTATCTCAAAACCTATGAGATGATTTGTTGAAAGTGTGCTTAATAA 501  
 QY 475 GATTACTATCTCAAAACCTATGAGATGATTTGTTGAAAGTGTGCTTAATAA 501  
 DB GATTACTATCTCAAAACCTATGAGATGATTTGTTGAAAGTGTGCTTAATAA 540  
 QY 481 CAAGAAAGGTGCAAGATGGCCAGACCTGATGTTATGCTGTGCTCAAGAGAGAG 540  
 DB CAAGAAAGGTGCAAGATGGCCAGACCTGATGTTATGCTGTGCTCAAGAGAGAG 501  
 QY 502 ----- 501  
 DB 502 ----- 501  
 QY 541 GCAAGCTCAGATGATCTGTTCTTCTTAAGAGAGCAATCTCCAAATACCTAATCTC 600  
 DB GCAAGCTCAGATGATCTGTTCTTCTTAAGAGAGCAATCTCCAAATACCTAATCTC 501  
 QY 502 ----- 501  
 DB 502 ----- 501  
 QY 601 ATATTAAAGCCAGGCTCTAGACAACTTATTACTTGTGATGCTGAGACTCCAC 660  
 DB ATATTAAAGCCAGGCTCTAGACAACTTATTACTTGTGATGCTGAGACTCCAC 552  
 QY 502 -----GCCAGGCTCTAGACAACTTATTACTTGTGATGCTGAGACTCCAC 552  
 DB -----GCCAGGCTCTAGACAACTTATTACTTGTGATGCTGAGACTCCAC 720  
 QY 661 CTGATAATTGCTCTTAATGCACTGCGTCTGTAATCCCAATPACTCTGGAACGTTCTAGT 720  
 DB CTGATAATTGCTCTTAATGCACTGCGTCTGTAATCCCAATPACTCTGGAACGTTCTAGT 612  
 QY 553 CTGATAATTGCTCTTAATGCACTGCGTCTGTAATCCCAATPACTCTGGAACGTTCTAGT 612  
 DB CTGATAATTGCTCTTAATGCACTGCGTCTGTAATCCCAATPACTCTGGAACGTTCTAGT 780  
 QY 721 GCCTGAGCTGTGTAAGTACAGGCGCAGCAAAAGTAACTTTCTTGGGAACTGGGT 780  
 DB GCCTGAGCTGTGTAAGTACAGGCGCAGCAAAAGTAACTTTCTTGGGAACTGGGT 672  
 QY 613 GCCTGAGCTGTGTAAGTACAGGCGCAGCAAAAGTAACTTTCTTGGGAACTGGGT 672

QY 781 AATGAGCAAAATACCTATGCGAACCATGATGCGCGGGCAGTAAATGCGAGCGATTGGAG 840  
 DB AATGAGCAAAATACCTATGCGAACCATGATGCGCGGGCAGTAAATGCGAGCGATTGGAG 732  
 QY 841 AAGGATTAACATCCAGCTGGAAGAGCCTGTTGACGCCCATCCGATTTATTCCAGAGCCAG 900  
 DB AAGGATTAACATCCAGCTGGAAGAGCCTGTTGACGCCCATCCGATTTATTCCAGAGCCAG 792  
 QY 733 AAGGATTAACATCCAGCTGGAAGAGCCTGTTGACGCCCATCCGATTTATTCCAGAGCCAG 792  
 DB AAGGATTAACATCCAGCTGGAAGAGCCTGTTGACGCCCATCCGATTTATTCCAGAGCCAG 960  
 QY 901 TTATATGCCCCCTAATTTGGCGCGCGGAGGAAGAATGTATGCGCCCTCCAGATGATTC 960  
 DB TTATATGCCCCCTAATTTGGCGCGCGGAGGAAGAATGTATGCGCCCTCCAGATGATTC 852  
 QY 793 TTATATGCCCCCTAATTTGGCGCGCGGAGGAAGAATGTATGCGCCCTCCAGATGATTC 852  
 DB TTATATGCCCCCTAATTTGGCGCGCGGAGGAAGAATGTATGCGCCCTCCAGATGATTC 1020  
 QY 961 ATGAAAGTGCAGAGAAATACAGTAGATGCAATGACGTTACCTGCGCAATGCTACATTTGATG 1020  
 DB ATGAAAGTGCAGAGAAATACAGTAGATGCAATGACGTTACCTGCGCAATGCTACATTTGATG 912  
 QY 853 ATGAAAGTGCAGAGAAATACAGTAGATGCAATGACGTTACCTGCGCAATGCTACATTTGATG 912  
 DB ATGAAAGTGCAGAGAAATACAGTAGATGCAATGACGTTACCTGCGCAATGCTACATTTGATG 1080  
 QY 1021 CGGCTGCTCAAGGTATGACACTTCTGAAAACCTGCTGTTAGACACACTCTCTGACAC 1080  
 DB CGGCTGCTCAAGGTATGACACTTCTGAAAACCTGCTGTTAGACACACTCTCTGACAC 972  
 QY 913 CGGCTGCTCAAGGTATGACACTTCTGAAAACCTGCTGTTAGACACACTCTCTGACAC 972  
 DB CGGCTGCTCAAGGTATGACACTTCTGAAAACCTGCTGTTAGACACACTCTCTGACAC 1032  
 QY 1081 ATTAGAAATTTGAGAAAGTGTAAATACATACACTCCAGAGAAAGATTTGCTTGA 1140  
 DB ATTAGAAATTTGAGAAAGTGTAAATACATACACTCCAGAGAAAGATTTGCTTGA 1032  
 QY 973 ATTAGAAATTTGAGAAAGTGTAAATACATACACTCCAGAGAAAGATTTGCTTGA 1032  
 DB ATTAGAAATTTGAGAAAGTGTAAATACATACACTCCAGAGAAAGATTTGCTTGA 1200  
 QY 1141 GGTGTGTGACCACTTCACTGAGAGGCAAAACATCTATCCGATTCTATGCTGACAGA 1200  
 DB GGTGTGTGACCACTTCACTGAGAGGCAAAACATCTATCCGATTCTATGCTGACAGA 1092  
 QY 1033 GGTGTGTGACCACTTCACTGAGAGGCAAAACATCTATCCGATTCTATGCTGACAGA 1092  
 DB GGTGTGTGACCACTTCACTGAGAGGCAAAACATCTATCCGATTCTATGCTGACAGA 1260  
 QY 1201 TTCTTATGTTGAAACCTTATGAAATGCTGGGCAATCAGGGCAATGATGCGATACGG 1260  
 DB TTCTTATGTTGAAACCTTATGAAATGCTGGGCAATCAGGGCAATGATGCGATACGG 1152  
 QY 1093 TTCTTATGTTGAAACCTTATGAAATGCTGGGCAATCAGGGCAATGATGCGATACGG 1152  
 DB TTCTTATGTTGAAACCTTATGAAATGCTGGGCAATCAGGGCAATGATGCGATACGG 1320  
 QY 1261 CACTCATTTTGTGACATGATACATACCTGCTGAGACCAAGATTTTAAACCATTAACA 1320  
 DB CACTCATTTTGTGACATGATACATACCTGCTGAGACCAAGATTTTAAACCATTAACA 1212  
 QY 1153 CACTCATTTTGTGACATGATACATACCTGCTGAGACCAAGATTTTAAACCATTAACA 1212  
 DB CACTCATTTTGTGACATGATACATACCTGCTGAGACCAAGATTTTAAACCATTAACA 1380  
 QY 1321 GACTACTGCTCTCTCTCTCTTACAAAGGCTGATGCGCCCAAGCTTGGCTGTCAT 1380  
 DB GACTACTGCTCTCTCTCTCTTACAAAGGCTGATGCGCCCAAGCTTGGCTGTCAT 1272  
 QY 1213 GACTACTGCTCTCTCTCTCTTACAAAGGCTGATGCGCCCAAGCTTGGCTGTCAT 1272  
 DB GACTACTGCTCTCTCTCTCTTACAAAGGCTGATGCGCCCAAGCTTGGCTGTCAT 1440  
 QY 1381 GTGCTGGGCTCCAGCGGAAGCAAGGCTGCGGCTGAGTATCCGGGCAAACTAAGAT 1440  
 DB GTGCTGGGCTCCAGCGGAAGCAAGGCTGCGGCTGAGTATCCGGGCAAACTAAGAT 1332  
 QY 1273 GTGCTGGGCTCCAGCGGAAGCAAGGCTGCGGCTGAGTATCCGGGCAAACTAAGAT 1332  
 DB GTGCTGGGCTCCAGCGGAAGCAAGGCTGCGGCTGAGTATCCGGGCAAACTAAGAT 1500  
 QY 1441 TATGCTACTGACAAACCAACCAACCAACCACTAGCTTGTGTGCTCACTTACTTTT 1500  
 DB TATGCTACTGACAAACCAACCAACCAACCACTAGCTTGTGTGCTCACTTACTTTT 1392  
 QY 1333 TATGCTACTGACAAACCAACCAACCAACCACTAGCTTGTGTGCTCACTTACTTTT 1392  
 DB TATGCTACTGACAAACCAACCAACCAACCACTAGCTTGTGTGCTCACTTACTTTT 1560  
 QY 1501 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAATCTCAGAGCAAG 1560  
 DB ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAATCTCAGAGCAAG 1452  
 QY 1393 ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAATCTCAGAGCAAG 1452  
 DB ATCATCAACTTGCATGATCAAGAAAGAAATCAAGCTGCTGGAATCTCAGAGCAAG 1620  
 QY 1561 CTGCTTCAACAGTACTGCTGACGCTTATGCGGAGGAGGCTTAAAGTCCAAAGTCA 1620  
 DB CTGCTTCAACAGTACTGCTGACGCTTATGCGGAGGAGGCTTAAAGTCCAAAGTCA 1512  
 QY 1453 CTGCTTCAACAGTACTGCTGACGCTTATGCGGAGGAGGCTTAAAGTCCAAAGTCA 1512  
 DB CTGCTTCAACAGTACTGCTGACGCTTATGCGGAGGAGGCTTAAAGTCCAAAGTCA 1680  
 QY 1621 CAACTGAATGGCCGACCTTATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 DB CAACTGAATGGCCGACCTTATGATGATGATGATGATGATGATGATGATGATGATG 1572  
 QY 1513 CAACTGAATGGCCGACCTTATGATGATGATGATGATGATGATGATGATGATGATG 1572  
 DB CAACTGAATGGCCGACCTTATGATGATGATGATGATGATGATGATGATGATGATG 1740  
 QY 1681 CGCCCGCTTGGGCGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATG 1740  
 DB CGCCCGCTTGGGCGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATG 1632  
 QY 1573 CGCCCGCTTGGGCGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATG 1632  
 DB CGCCCGCTTGGGCGCGGCGGAGCAATGATGATGATGATGATGATGATGATGATG 1779  
 QY 1741 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779  
 DB GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1671  
 QY 1633 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1671  
 DB GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG

RESULT 8  
 AF282886  
 LOCUS

AF282886

1922 bp

mRNA

linear

PRI 14-OCT-2000

DEFINITION	Homo sapiens heparanase-like protein HPA2b mRNA, complete cds.
ACCESSION	AJ282886
VERSION	AJ282886.1 GI:10801196
KEYWORDS	
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1922) McKenzie,E., Tyson,K., Stamps,A., Smith,P., Turner,P., Barry,R., Hitcock,M., Patel,S., Barry,E., Stubberfield,C., Terrett,J. and Page,M. Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member Biochem. Biophys. Res. Commun. 276 (3), 1170-1177 (2000).
TITLE	
JOURNAL	
MEDLINE	20483645
PUBMED	11027606
REFERENCE	2 (bases 1 to 1922) McKenzie,E.A., Tyson,K. and Stamps,A. Direct Submission Submitted (28-JUN-2000) Biology Group, Oxford Glycosciences, Quadrant, Abingdon Science Park, Abingdon OX14 3YS, England Location/Qualifiers 1..1922
FEATURES	
SOURCE	

ORIGIN

Query Match	79.7%;	Score 1417.8;	DB 9;	Length 1922;
Best Local Similarity	90.1%;	Pred. No. 0;		
Matches 1603;	Conservative	0;	Mismatches 2;	Indels 174;
				Gaps 1;

QY	1	ATGAGGAGCTTTGTGACCTTCCCTGAAACATATGCTTCACAACATCCGAGCCCGCGCG	60
Db	61	ATGAGGAGCTTTGTGACCTTCCCTGAAACATGCTTCACAACATCCGAGCCCGCGCG	120
QY	61	TGCCATGAGCCCGGGGGCTCTCTACTTGGGCTCTGTGGCTCAGATCTTCCCTTCCGCCAG	120
Db	121	TGCCATGAGCCCGGGGGCTCTCTACTTGGGCTCTGTGGCTCAGATCTTCCCTTCCGCCAG	180
QY	121	GCTGAGAGCAGAGAGACCTTGGCTGTGAGACAGAGCTCAGAGTTTGAGAGAAAAGACCTG	180
Db	181	GCTGAGAGCAGAGAGACCTTGGCTGTGAGACAGAGCTCAGAGTTTGAGAGAAAAGACCTG	240
QY	181	ATTCTACTGTATGTAGACACCMAGAACCGAGTCAGAGACAGTCAATGAGAACTTCTCTCT	240
Db	241	ATTCTACTGTATGTAGACACCMAGAACCGAGTCAGAGACAGTCAATGAGAACTTCTCTCT	300
QY	241	CTGAGAGCTGGAACCGCTGCATCATTTATATATGCTGTGGCTCGATTTCTTAAGCTCCAAAGGC	300
Db	301	CTGAGAGCTGGAACCGCTGCATCATTTATATATGCTGTGGCTCGATTTCTTAAGCTCCAAAGGC	360
QY	301	TTGTGACCCCTGGCCCGAGGACTTTTCGCCCCCTTTTCTGCGCTTGGGGGCAAAAGAAC	360

Db	361	TTGGTGAACCTGGCCCGGGGACCTTTGGCCCGCCCTTCTGGGCTTCGGGGGCAAAAGGACC	420
Oy	361	GACTTCTGAGTTCGCAAGACTGAGAGAACCCGCGCAAAAAGCCGGGGGACCCGGGGCCG	420
Db	421	GACTTCTGAGTTCGCAAGACTGAGAGAACCCGCGCAAAAAGCCGGGGGACCCGGGGCCG	480
Oy	421	GATTACTATCTCAAAACTATGAGATGACATTGTTGAAAGTATGTTGCCCTTAGATAA	480
Db	481	GATTACTATCTCAAAACTATGAGATGACATTGTTGAAAGTATGTTGCCCTTAGATAA	540
Oy	481	CAGAAAGGCTGCAGAGATTGCCGAGACCCCTGATGTTATGCTGGTGCTCAAAAGGAGAA	540
Db	541	CAGAAAGGCTGCAGAGATTGCCGAGACCCCTGATGTTATGCTGGTGCTCAAAAGGAGAA	600
Oy	541	GCAGCTCAGATGCATCTGGTCTCTTAAAGAGCAATTTCCCAATCTTACAGTAATCTC	600
Db	601	GCAGCTCAGATGCATCTGGTCTCTTAAAGAGCAATTTCCCAATCTTACAGTAATCTC	660
Oy	601	ATATTACAGACAGAGTCTTAGACAAACTTATTAATCTTGTGATTTGCTGCATCCAC	660
Db	661	ATAATTACA-----	669
Oy	661	CTGATATTGGCTCTAAATGCACTGGCTGTATTCCTCAATATCTCTGGAACAGTTCTAGT	720
Db	670	-----	669
Oy	721	GCCTGTAGTGTGTAAGTACAGGCGCAGCAAAAAGTACAACTATTTCTGGAACTGGGT	780
Db	670	-----	669
Oy	781	AATGAGCCAAATTAATCTATCGAACCATGATGGCCGGGAGTAATGGCAGCCAGTTGGGA	840
Db	670	---GAGCAATATATCTATCGAACCATGATGGCCGGGAGTAATGGCAGCCAGTTGGGA	726
Oy	841	AAGGATTAACATCCAGCTGGAAGAGGCTTTGCAAGCCATCCGGAATTATTTTCAGAGCCAG	900
Db	727	AAGGATTAACATCCAGCTGGAAGAGGCTTTGCAAGCCATCCGGAATTATTTTCAGAGCCAG	786
Oy	901	TTATATGGCCCTAATATTGGGCGGGCGAGAGAAATGTCAATGCCCTCTTAGATGATTC	960
Db	787	TTATATGGCCCTAATATTGGGCGGGCGAGAGAAATGTCAATGCCCTCTTAGATGATTC	846
Oy	961	ATGAAGGTGCGAGAAATGACATGATGACATGACATCTGCGAACAATTGCTACATGATGAC	1020
Db	847	ATGAAGGTGCGAGAAATGACATGATGACATGACATCTGCGAACAATTGCTACATGATGAC	906
Oy	1021	CGGGTGGTCAAGGTGATGGAATCTCTGAAAATCTGCCCTGTTAGACACAATCTCTGACAG	1080
Db	907	CGGGTGGTCAAGGTGATGGAATCTCTGAAAATCTGCCCTGTTAGACACAATCTCTGACAG	966
Oy	1081	ATTAGGAAAAATTCAAGAAAGTGTTAATATACATCACTCCAGAAAAGAAATTTGGCTTGAA	1140
Db	967	ATTAGGAAAAATTCAAGAAAGTGTTAATATACATCACTCCAGAAAAGAAATTTGGCTTGAA	1026
Oy	1141	GGTGTGGTGAACACTCACTGAGCTGAGGAGCAAAACAATTAATCCATTTCTATGCTGCAGGA	1200
Db	1027	GGTGTGGTGAACACTCACTGAGCTGAGGAGCAAAACAATTAATCCATTTCTATGCTGCAGGA	1086
Oy	1201	TTCTTATGGTTGAACAATTAGAGATCTGGCCATCAGGGCATTTGATGTCGTATACGG	1260
Db	1087	TTCTTATGGTTGAACAATTAGAGATCTGGCCATCAGGGCATTTGATGTCGTATACGG	1146
Oy	1261	CACGATTTTTTTGACATGATATACATCACTGCTGGAACGAAATTTTAAACCCATTAACA	1320
Db	1147	CACGATTTTTTTGACATGATATACATCACTGCTGGAACGAAATTTTAAACCCATTAACA	1206
Oy	1321	GACTACTGGCTCTCTCTCTCTCAACAAGCGCTGATCGGCCCCCAAGCTTTGGCTGTGCAT	1380
Db	1207	GACTACTGGCTCTCTCTCTCTCAACAAGCGCTGATCGGCCCCCAAGCTTTGGCTGTGCAT	1266
Oy	1381	GTTGGCTGGGCTCCAGCGGAAGCCAGCGGCTGGCCGAGATATCCGGGACAAACTAGAGATT	1440
Db	1267	GTTGGCTGGGCTCCAGCGGAAGCCAGCGGCTGGCCGAGATATCCGGGACAAACTAGAGATT	1326

Qy	1444	TATGTCAGTCGCAAAACACAAACAAACACAACTACAGTCGTGGTGCATTAACCTTTT	1500
Db	1327	TATGTCAGTCGCAAAACACAAACACAACTACAGTCGTGGTGCATTAACCTTTT	1386
Qy	1501	ATCATCAACTTGCATTCGATCAGAAAGAAATCAAGTCGTGGTGCATTCACAGACAAAG	1560
Db	1387	ATCATCAACTTGCATTCGATCAGAAAGAAATCAAGTCGTGGTGCATTCACAGACAAAG	1446
Qy	1561	CTGCTTCAACAGTACCTGCTGTGACAGCCCTATGGGACGAGAGGGCTTAAAGTCCAAATCAGTG	1620
Db	1447	CTGCTTCAACAGTACCTGCTGTGACAGCCCTATGGGACGAGAGGGCTTAAAGTCCAAATCAGTG	1506
Qy	1621	CAACTGAAATGGCCAGCCCTTAGTGTATGGTGTGACGACGGGAAACCTCCACGAATTGAAGCCC	1680
Db	1507	CAACTGAAATGGCCAGCCCTTAGTGTATGGTGTGACGACGGGAAACCTCCACGAATTGAAGCCC	1566
Qy	1681	CGCCCCCTTCGGGGCGGGCGGACATTTGGTTCATCCCTCCAGTCACACATGGGCTTTTATGTG	1740
Db	1567	CGCCCCCTTCGGGGCGGGCGGACATTTGGTTCATCCCTCCAGTCACACATGGGCTTTTATGTG	1626
Qy	1741	GTCAGGAATGTCAATGCTTTGGCCCTGGCCGCTACCGGATTA	1779
Db	1627	GTCAGGAATGTCAATGCTTTGGCCCTGGCCGCTACCGGATTA	1665

RESULT 9			
AX286343			
LOCUS	AX286343	2326 bp	DNA
DEFINITION	Sequence 1 from Patent WO0181569.		linear
ACCESSION	AX286343		
VERSION	AX286343.1	GI:17048588	

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Heinrikson, R.L. and Bienkowski, M.J.
TITLE	Hepatarnase II, a human hepatarnase paralog
JOURNAL	Patent: WO 0181569-A 1 01-NOV-2001; Pharmacia & Upjohn Company (US)
FEATURES	Location/Qualifiers
SOURCE	1..2326
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"

Query Match      79.7%;    Score 1417.8;    DB 6;    Length 2326;

QY 1 ATGAGAGGTCCTTGGAGCCTCCCTGTAAGCAATGCCCTCAGAACTCCGGCCCCGGCG 60

Db 25 ATGAGAGGTCCTTGGAGCCTCCCTGTAAGCAATGCCCTCAGAACTCCGGCCCCGGCG 84

QY 61 TGCCTAGCCCCGGAGGAGCTCTACTTGGCTGTGTTCATCTCTCCCTTCTCCAG 120

Db 85 TGCCTAGCCCCGGAGGAGCTCTACTTGGCTGTGTTCATCTCTCCCTTCTCCAG 144

QY 121 GGTGAGACAGGAGACCTTGCCTGTAGACAGAGCTGACAGGTTTAAAGAAAAGACCTG 180

Db 145 GGTGAGACAGGAGACCTTGCCTGTAGACAGAGCTGACAGGTTTAAAGAAAAGACCTG 204

QY 181 ATTCTACTGATGTGAGACCAAGAACCCAGTCAGAGACGTCAATGAGAACTTCTCTCT 240

Db 205 ATTCTACTGATGTGAGACCAAGAACCCAGTCAGAGACGTCAATGAGAACTTCTCTCT 264

QY 241 CTGCAGCTGGATCCGTCATCATTTATATATGCTGGCTCGATTTTCTCTAAGCTCAAAGGC 300

Db 265 CTGCAGCTGGATCCGTCATCATTTATATATGCTGGCTCGATTTTCTCTAAGCTCAAAGGC 324

[illegible]



Db 1231 GTGGCTGGGCTCCAGCGGAGCCAGCGCTGGCCGAGTATCCCGGACAAACTAGAGATT 1290  
 Qy 1441 TATGCTCACTGACAAACCAACCAACCAACCAACTAGCTTGGTGGTCACTTACCTTTT 1500  
 Db 1291 TATGCTCACTGACAAACCAACCAACCAACCAACTAGCTTGGTGGTCACTTACCTTTT 1350  
 Qy 1501 ATCATCACTGACAAACCAACCAACCAACCAACTAGCTTGGTGGTCACTTACCTTTT 1560  
 Db 1351 ATCATCACTGACAAACCAACCAACCAACCAACTAGCTTGGTGGTCACTTACCTTTT 1410  
 Qy 1561 CTGGTTCACCAAGTACCTCTGAGCCCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
 Db 1411 CTGGTTCACCAAGTACCTCTGAGCCCTATGAGGAGAGAGAGAGAGAGAGAGAGAG 1470  
 Qy 1621 CAAGTGAATGGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 Db 1471 CAAGTGAATGGCCAGCCCTTATGATGATGATGATGATGATGATGATGATGATGAT 1530  
 Qy 1681 CGCCCTCTGAGGAG 1740  
 Db 1531 CGCCCTCTGAGGAG 1590  
 Qy 1741 GTCAAGATGTCATAGCTTTGGCTGGCCGCTACCGATTA 1779  
 Db 1591 GTCAAGATGTCATAGCTTTGGCTGGCCGCTACCGATTA 1629

## RESULT 10

AX180832 2462 bp DNA linear PAT 06-AUG-2001  
 LOCUS AX180832 Sequence 3 from Patent WO0146392.  
 DEFINITION AX180832  
 ACCESSION AX180832.1 GI:15132668  
 VERSION AX180832.1  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Mckenzie, E.A., Stamps, A.C., Terrett, J.A. and Tyson, K.L.  
 Homologues of human heparanase and splice variants thereof  
 Patent: WO 0146392-A 3 28-JUN-2001;  
 JOURNAL Oxford Glycosciences (UK) Limited (GB)  
 FEATURES  
 source  
 1..2462  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 601..2205  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC5009.1"  
 /db\_xref="GI:15132668"  
 /translation="MRVLCAPEAPSSNSRPPACLAIPALYALILHLSLSSQADR  
 RPLPDRAPAGLKEKTLILDVSTKNPVTAVENFLSLQDPSIIHDMWLFSLSKRLV  
 TLARGLSPALRFGKRTDFLQFNLRNPARKRGPGPDYILKRVEDIVRSVDALDK  
 OKGKIACHPVMLELOREKAAOMHLVLEKEQFNTYSNLILTEPNRYTHGSAVWG  
 SOLKDYIOLKSLQPIRISRAISLIGNIGRPKNRYIALIDGPMKAVAGTVAWVG  
 HCYIDGRVAVKMDPLKTRLDLIDSLDOILOKVVNTYTPKKIILBQVTTAGAGTN  
 LSGDFAAFPLWNTLGMANGSIDIVYIHSFEDGYNHNDQNNPILPDVLSILYLR  
 LIGPVLAVHAGLQKRPGRGIVRDKRTIYACTNHNHNTYVGSITLILNHRK  
 KIKLAGTIRKLVHGYLLQPYGQGLKSKSVQNLGQPLVWVDGTLPELPRLRAG  
 RLVLPVTMGFFVKNVNALACRR"  
 CDS  
 601..2205  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC5009.1"  
 /db\_xref="GI:15132668"  
 /translation="MRVLCAPEAPSSNSRPPACLAIPALYALILHLSLSSQADR  
 RPLPDRAPAGLKEKTLILDVSTKNPVTAVENFLSLQDPSIIHDMWLFSLSKRLV  
 TLARGLSPALRFGKRTDFLQFNLRNPARKRGPGPDYILKRVEDIVRSVDALDK  
 OKGKIACHPVMLELOREKAAOMHLVLEKEQFNTYSNLILTEPNRYTHGSAVWG  
 SOLKDYIOLKSLQPIRISRAISLIGNIGRPKNRYIALIDGPMKAVAGTVAWVG  
 HCYIDGRVAVKMDPLKTRLDLIDSLDOILOKVVNTYTPKKIILBQVTTAGAGTN  
 LSGDFAAFPLWNTLGMANGSIDIVYIHSFEDGYNHNDQNNPILPDVLSILYLR  
 LIGPVLAVHAGLQKRPGRGIVRDKRTIYACTNHNHNTYVGSITLILNHRK  
 KIKLAGTIRKLVHGYLLQPYGQGLKSKSVQNLGQPLVWVDGTLPELPRLRAG  
 RLVLPVTMGFFVKNVNALACRR"  
 ORIGIN  
 Query Match 79.7%; Score 1417.8; DB 6; Length 2462;  
 Best Local Similarity 90.1%; Pred. No. 0; Mismatches 1;  
 Matches 1603; Conservative 0; Indels 174; Gaps 1;  
 Qy 1 ATGAGGGTCTTGTGCTTCCCTGAGCCATGCGCTTCAGCAACTCCGCGCGCGG 60

Qy 1 ATGAGGGTCTTGTGCTTCCCTGAGCCATGCGCTTCAGCAACTCCGCGCGCGG 60

Db 601 ATGAGGGTCTTGTGCTTCCCTGAGCCATGCGCTTCAGCAACTCCGCGCGCGG 660  
 Qy 61 TGCCTAAGCCCCGGGGGCTCTCTACTTGGCTCTGTGTGCTCACTCTCCCTTTCTCCAG 120  
 Db 661 TGCCTAAGCCCCGGGGGCTCTCTACTTGGCTCTGTGTGCTCACTCTCCCTTTCTCCAG 720  
 Qy 121 GCTGAGACAGAGAGACCTTGTGCTGTAGACAGAGCTGACAGGTTTGAAGAAAAGACCTG 180  
 Db 721 GCTGAGACAGAGAGACCTTGTGCTGTAGACAGAGCTGACAGGTTTGAAGAAAAGACCTG 780  
 Qy 181 ATTCTACTGATGTGACACCAAGAACCCAGTCAAGATGATGATGATGATGATGATGATGAT 240  
 Db 781 ATTCTACTGATGTGACACCAAGAACCCAGTCAAGATGATGATGATGATGATGATGATGAT 840  
 Qy 241 CTGCAAGCTGATCCGTCATCATTCATGATGATGATGATGATGATGATGATGATGATGAT 300  
 Db 841 CTGCAAGCTGATCCGTCATCATTCATGATGATGATGATGATGATGATGATGATGATGAT 900  
 Qy 301 TTGGTGAACCTTGGCCCGGGAGCTTTTGGCCCGCTTTCTGCGCTTGGGGGCAAGAGACC 360  
 Db 901 TTGGTGAACCTTGGCCCGGGAGCTTTTGGCCCGCTTTCTGCGCTTGGGGGCAAGAGACC 960  
 Qy 361 GACTTCCTGACAGTCCAGAACCTGAGAAACCCGGGCAAAAGCCGGGGGCGCGGCGCG 420  
 Db 961 GACTTCCTGACAGTCCAGAACCTGAGAAACCCGGGCAAAAGCCGGGGGCGCGGCGCG 1020  
 Qy 421 GATTACTATCTCAAAAACCTATGAGAGATGACATTTGTTGAAATGATGTTGCTTAGATPAA 480  
 Db 1021 GATTACTATCTCAAAAACCTATGAGAGATGACATTTGTTGAAATGATGTTGCTTAGATPAA 1080  
 Qy 481 CAGAAAGCTGAGAGATTGCGCAGACCCCTGATGTTAGTCTGCTGCTCCAAAGGAGAGAG 540  
 Db 1081 CAGAAAGCTGAGAGATTGCGCAGACCCCTGATGTTAGTCTGCTGCTCCAAAGGAGAGAG 1140  
 Qy 541 GCAGCTCAGATCATCTGCTTCTTCTTAAAGAGACATTTCCCATATCTTACAGTATCTC 600  
 Db 1141 GCAGCTCAGATCATCTGCTTCTTCTTAAAGAGACATTTCCCATATCTTACAGTATCTC 1200  
 Qy 601 AATTATACGCAAGGCTCTTACAAACCTTTTAACTTTGCTGATGCTGAGATCCAC 660  
 Db 1201 AATTATACGCAAGGCTCTTACAAACCTTTTAACTTTGCTGATGCTGAGATCCAC 1209  
 Qy 661 CTGATATTTGCTCTAATGACACTGCGTGTATCCCAATTAACCTCGGACAGATTCTAGT 720  
 Db 1210 CTGATATTTGCTCTAATGACACTGCGTGTATCCCAATTAACCTCGGACAGATTCTAGT 1209  
 Qy 721 GCCCTGAGTCTGTGAAGTACAGCCGCCAGCAAAAGTACACATTTCTGGGAAGTGGGT 780  
 Db 1210 GCCCTGAGTCTGTGAAGTACAGCCGCCAGCAAAAGTACACATTTCTGGGAAGTGGGT 1209  
 Qy 781 AATGAGCCAAATTAATATGACATGACATGATGCGCGGAGTAAATGAGCAGCCAGTTGGGA 840  
 Db 1210 AATGAGCCAAATTAATATGACATGACATGATGCGCGGAGTAAATGAGCAGCCAGTTGGGA 1266  
 Qy 841 AAGGATTAATATGACATGACATGATGCGCGGAGTAAATGAGCAGCCAGTTGGGA 900  
 Db 1267 AAGGATTAATATGACATGACATGATGCGCGGAGTAAATGAGCAGCCAGTTGGGA 1326  
 Qy 901 TTAATGCGCTTAATATGAGCGCGGAGTAAATGAGCAGCCAGTTGGGA 960  
 Db 1327 TTAATGCGCTTAATATGAGCGCGGAGTAAATGAGCAGCCAGTTGGGA 1386  
 Qy 961 ATGAGAGTGGCAGAGATGACATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1020  
 Db 1387 ATGAGAGTGGCAGAGATGACATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1446  
 Qy 1021 CGGAGTGGTCAAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1080  
 Db 1447 CGGAGTGGTCAAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1506  
 Qy 1081 ATTAGAATATGAGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1140  
 Db 1507 ATTAGAATATGAGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1566







D	b		415	GACTTCCTGCAAGTTCCAGAACCTGAGGACC	CGCGAAAAGCCGCGGGGGCCCGGGCCG	474
O	y		421	GATTACTATCTCAAAAACCTATGAGATGACAT	TTGTTGAAAGTATGTTGGCTTATGATATAA	480
D	b		475	GATTACTATCTCAAAAACCTATGAGATGACAT	TTTCCGAGTATGTTGGCTTATGATATAA	534
O	y		481	CAGAAAAGCTGCAAGATTGGCCACACCCCTG	ATGTATGCTGTGATGCTCCAAAGGAGAG	540
D	b		535	CAGAAAAGCTGCAAGATTGGCCACACCCCTG	ATGTATGCTGTGATGCTCCAAAGGAGAG	594
O	y		541	GCAGCTCAGATGCATCTGGCTTCTCTAAAG	AGCAATTCCTCAATCTTCAAGTAATCTC	600
D	b		595	GCAGCTCAGATGCATCTGGCTTCTCTAAAG	AGCAATTCCTCAATCTTCAAGTAATCTC	654
O	y		601	ATATTAACAGCCAGAGTCTCTAGACAAACT	TTTATTAATCTTGTGATGCTCTGCACTCAC	660
D	b		655	ATATTAACA		663
O	y		661	CTGATATTGTGCTTAATATGCACTGCGTCG	TAAATCCCAATAACTCCTTGAAACAGTTCT	720
D	b		664			663
O	y		721	GCCTTGAGTCTGTTGAAGTACAGCGCAGCA	AAAAAGTACACATTTCTTGGGAATGGAGT	780
D	b		664			663
O	y		781	AATGAGCCAAATTAATCTATCCGACCATG	ATGCGCGGAGTAATATGGCAGCAGTGGGA	840
D	b		664	--GAGCCAAATTAATCTATCCGACCATG	ATGCGCGGAGTAATATGGCAGCAGTGGGA	720
O	y		841	AAGATTACATCCAGCTGAAAGGCTGTGGAC	CCCATCCGAGATTATTCGAGACCGAC	900
D	b		721	AAGATTACATCCAGCTGAAAGGCTGTGGAC	CCCATCCGAGATTATTCGAGACCGAC	780
O	y		901	TTATATGGCCCTTAATATTGGCGCGCGCA	AGGAAGATGCATCGCCCTCTAGATGCAATTC	960
D	b		781	TTATATGGCCCTTAATATTGGCGCGCGCA	AGGAAGATGCATCGCCCTCTAGATGCAATTC	840
O	y		961	ATGAAGGTGGCAGGAAGTACAGTAGATGAT	CACTTCCGCAACAATTTGGCTTGAA	1020
D	b		841	ATGAAGGTGGCAGGAAGTACAGTAGATGAT	CACTTCCGCAACAATTTGGCTTGAA	900
O	y		1021	CGGGTGGTCAAGGTGATGAGACCTTCCGAA	ACTGGCTGTAGACACACTCTGACCGAG	1080
D	b		901	CGGGTGGTCAAGGTGATGAGACCTTCCGAA	ACTGGCTGTAGACACACTCTGACCGAG	960
O	y		1081	ATTAGGAATATTCAGAAAAGGTATATATCA	TAACTCCAGAAAAGAAATTTGGCTTGAA	1140
D	b		961	ATTAGGAATATTCAGAAAAGGTATATATCA	TAACTCCAGAAAAGAAATTTGGCTTGAA	1020
O	y		1141	GGTGGTGGACCACTCAGCTGAGGAGCAAA	CAATCTATCCGATTCCTATGCTGCAGGA	1200
D	b		1021	GGTGGTGGACCACTCAGCTGAGGAGCAAA	CAATCTATCCGATTCCTATGCTGCAGGA	1080
O	y		1201	TTCTTATGTTGAACACTTTAGAAATGCTGG	CCATTCAGGCAATGATGTCGTATACG	1260
D	b		1081	TTCTTATGTTGAACACTTTAGAAATGCTGG	CCATTCAGGCAATGATGTCGTATACG	1140
O	y		1261	CACATCATTTTGAACAAGATACATACCTG	AGCAAGAAATTTTAAACCATTAACA	1320
D	b		1141	CACATCATTTTGAACAAGATACATACCTG	AGCAAGAAATTTTAAACCATTAACA	1200
O	y		1321	GACTATCTGGCTCTCTCTCTCTATCAAGG	CGCTGATCGGACCCCAAGTCTTGCTGTGAT	1380
D	b		1201	GACTATCTGGCTCTCTCTCTCTCTATCAAG	CGCTGATCGGACCCCAAGTCTTGCTGTGAT	1260
O	y		1381	GTGGCTGGGGCTCCAGCGGAGCAACAGGCT	GGAGGATCCGGGACAAACTATAGAGATT	1440
D	b		1261	GTGGCTGGGGCTCCAGCGGAGCAACAGGCT	GGAGGATCCGGGACAAACTATAGAGATT	1320
O	y		1441	TATGTCACTGCAAAACCAACCAACAATA	GTGTTGTTGGTGCATTACAATTTT	1500

Df		1321	TATCTCACTGCGACAAACCCACCAACAACAACAAGTTCGTGGTGTCATTACCTTTT	1380
OY		1501	ATCATCAACTGTGCATGATCAGAAGAAAATAAAGTGGCTGGGACTTCAGAGCAAG	1560
Df		1381	ATCATCAACTGTGCATGATCAGAAGAAAATAAAGTGGCTGGGACTTCAGAGCAAG	1440
OY		1561	CTGGTTCACCAAGTACTGCTGCAAGCCCTATGGGAGGAGGAGGCTTAAGTCCAAGTCAGTG	1620
Df		1441	CTGGTTCACCAAGTACTGCTGCAAGCCCTATGGGAGGAGGAGGCTTAAGTCCAAGTCAGTG	1500
OY		1621	CAACTGAATGGCGCACGCCCTTAGTATGATGATGAGAGCGAGGAGCCCTCCAGAAATTGAAGCCC	1680
Df		1501	CACCTGAATGGCGCACGCCCTTAGTATGATGATGAGAGCGAGGAGCCCTCCAGAAATTGAAGCCC	1560
OY		1681	CGCCCCCTTGGGAGCGGAGCCGAGACATTTGTCATCCCTCCAGTCAACATAGGAGCTTTATATGTG	1740
Df		1561	CGCCCCCTTGGGAGCGGAGCCGAGACATTTGTCATCCCTCCAGTCAACATAGGAGCTTTATATGTG	1620
OY		1741	GTCAGAAATGTCATAGCTTTGGGAGCTGGCGGATCCAGATTA	1779
Df		1621	GTCAGAAATGTCATAGCTTTGGGAGCTGGCGGATCCAGATTA	1659
RESULT 13				
AX190449				
LOCUS	AX190449	3943 bp	DNA	linear PAT 08-AUG-2001
DEFINITION	Sequence 1 from Patent WO0148161.			
ACCESION	AX190449			
VERSION	AX190449.1	GI:15143802		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euleleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
JOURNAL	Human heparanase-related polypeptide and nucleic acid			
FEATURES	Patent: WO 0148161-A 1 05-JUL-2001;			
source	SCHERING AKTIEGENSIEDLSCHAFT (DE)			
	Location/Qualifiers			
	1..3943			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	72.5%; Score 1289.8; DB 6; Length 3943;			
Best Local Similarity	89.2%; Pred. No. 0;			
Matches 1475; Conservative	1; Mismatches	3; Indels	174; Gaps	1;
OY		127	GACAGAGAACCTTGCCTGTGAGACAGAGCTGCAAGTTTGAAGAAAAAGACCTGATTCTA	186
Df		1	GACAGAGAACCTTGCCTGTGAGACAGAGCTGCAAGTTTGAAGAAAAAGACCTGATTCTA	60
OY		187	CTGATGTGAGACACCAAGAACCCGATGACAGACATCATAGAACTTCTCTCTGAC	246
Df		61	CTGATGTGAGACACCAAGAACCCGATGACAGACATCATAGAACTTCTCTCTGAC	120
OY		247	CTGATCCGTCATCATTCATGATGAGGCTGAGCTCATTTTCTTAAGCTCCAAACGTTGGTG	306
Df		121	CTGATCCGTCATCATTCATGATGAGGCTGAGCTCATTTTCTTAAGCTCCAAACGTTGGTG	180
OY		307	AACCTGGCCGGGGAATTTGGCCCGCTTTTGTGGCTTTGGGGGCAAAAGAACCGAATTC	366
Df		181	AACCTGGCCGGGGAATTTGGCCCGCTTTTGTGGCTTTGGGGGCAAAAGAACCGAATTC	240
OY		367	CTGCAATTCAGAAACCTGAGGAAACCGGCGAAAAAGCCGGGGGCCCGGGCCCGGATTAAC	426
Df		241	CTGCAATTCAGAAACCTGAGGAAACCGGCGAAAAAGCCGGGGGCCCGGGCCCGGATTAAC	300
OY		427	TATCTCAAAAATCTATGAGAGATGATTTTTCAGATGATGTTGCTTATGATTAACAGAA	486
Df		301	TATCTCAAAAATCTATGAGAGATGATTTTTCAGATGATGTTGCTTATGATTAACAGAA	360

```

QY 487 GGCTGAAGATGCGCCAGACCCCTGATGTTATGCTGCTGCTCCAAAGGAGAGGAGCT 546
DB 361 GGCTGAAGATGCGCCAGACCCCTGATGTTATGCTGAGAGCTCCAAAGGAGAGGAGCT 420
QY 547 CAGATGATCTGCTGCTTCTTAAAGAGCAATCTCCAAATCTTAAGTAATCTCATATTA 606
DB 421 CAGATGATCTGCTGCTTCTTAAAGAGCAATCTCCAAATCTTAAGTAATCTCATATTA 480
QY 607 ACAGCAGGCTCTAGACAAACTTATTAAGCTTGTGATGCTGTGAGCCACCTGATA 666
DB 481 ACA----- 483
QY 667 TTGCTCTAAATGCACTGCGCTGTAAATCCAAATTAATCTGGAACAGTTCTAGTCCCTG 726
DB 484 ----- 483
QY 727 AGTCTGTGAAGTACAGGCGCAGCAAAAGTACAACATTTCTTGGAACCTGGTAATAG 786
DB 484 -----GAG 486
QY 787 CCAATTAATCTGCGACCATGATGCGCGGAGTAATGCGACGCTTGGAAGAGAT 846
DB 487 CCAATTAATCTGCGACCATGATGCGCGGAGTAATGCGACGCTTGGAAGAGAT 546
QY 847 TACATCCAGCTGAAGAGCTTGTGAGCCCATCCGGAATTTATTCAGAGCCAGCTTATAT 906
DB 547 TACATCCAGCTGAAGAGCTTGTGAGCCCATCCGGAATTTATTCAGAGCCAGCTTATAT 606
QY 907 GAGCCATTAATGCGCGGCGGAGAGATGTATCCGCTCTGATGATGATCATGAAG 966
DB 607 GAGCCATTAATGCGCGGCGGAGAGATGTATCCGCTCTGATGATGATCATGAAG 666
QY 967 GTGCGAGAGTACAGTATGATGCACTTCTGCGCAATTCATTCATTTATGCGCGGCTG 1026
DB 667 GTGCGAGAGTACAGTATGATGCACTTCTGCGCAATTCATTCATTTATGCGCGGCTG 726
QY 1027 GTCAAGGTGATGAGCTTCTGTAAGAACTGCGCTGTTAGACACCTCTGACAGATTAG 1086
DB 727 GTCAAGGTGATGAGCTTCTGTAAGAACTGCGCTGTTAGACACCTCTGACAGATTAG 786
QY 1087 AAAATTCAAGAAAGTGTATATATACATCACTTCAGAGAAAGATTGGCTTGAAGCTGTG 1146
DB 787 AAAATTCAAGAAAGTGTATATATACATCACTTCAGAGAAAGATTGGCTTGAAGCTGTG 846
QY 1147 GTGACCACTCAGCTGAGGAGGAGCAAAACAATCTATCCGATTCCTATGCTGAGAGATCTTA 1206
DB 847 GTGACCACTCAGCTGAGGAGGAGCAAAACAATCTATCCGATTCCTATGCTGAGAGATCTTA 906
QY 1207 TGGTTGAACACTTTAGGAATGCTGGCAATCAGGGCAATTGATGCTGATACGGCACTCA 1266
DB 907 TGGTTGAACACTTTAGGAATGCTGGCAATCAGGGCAATTGATGCTGATACGGCACTCA 966
QY 1267 TTTTGTGACATGATATCAATCACTCTGTGACAGAAATTTTAAACCATTAACAAGATAC 1326
DB 967 TTTTGTGACATGATATCAATCACTCTGTGACAGAAATTTTAAACCATTAACAAGATAC 1026
QY 1327 TGGCTCTCTCTCTCTCAAGCGCTGATCGGCGCCCAAGTCTTGTGCTGTGCTGCT 1386
DB 1027 TGGCTCTCTCTCTCTCAAGCGCTGATCGGCGCCCAAGTCTTGTGCTGTGCTGCT 1086
QY 1387 GGGCTCCAGCGGAGGAGCCGCGCTGCGGAGTATCCGGAGCAAACTTAAGATTTATGCT 1446
DB 1087 GGGCTCCAGCGGAGGAGCCGCGCTGCGGAGTATCCGGAGCAAACTTAAGATTTATGCT 1146
QY 1447 CACTGCAAAACCAACAACAACAACAATCACTGCTGCTGCTCAATTAATTTATATC 1506
DB 1147 CACTGCAAAACCAACAACAACAACAACAATCACTGCTGCTGCTCAATTAATTTATATC 1206
QY 1507 AACTTGATGATCAAGAAAGAAATCAAGCTGCTGCTGCTCAAGAGCAAGCTGCT 1566
DB 1207 AACTTGATGATCAAGAAAGAAATCAAGCTGCTGCTGCTCAAGAGCAAGCTGCT 1266

```

```

QY 1567 CACAGTACTGCTGCTCAGACCCCTATGGGAGAGGGGCTTAAAGTCCAGTCAAGTCACTG 1626
DB 1267 CACAGTACTGCTGCTCAGACCCCTATGGGAGAGGGGCTTAAAGTCCAGTCAAGTCACTG 1326
QY 1627 AATGCGCAGCCCTTATGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
DB 1327 AATGCGCAGCCCTTATGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
QY 1687 CTTGCGGCGCGGCGGAGCAATGGTCAATCCCTCAGTCAACATAGGCTTTATGCTGAG 1746
DB 1387 CTTGCGGCGCGGCGGAGCAATGGTCAATCCCTCAGTCAACATAGGCTTTATGCTGAG 1446
QY 1747 AATGCAATGCTTTGGCTGCGCTGACCGATTA 1779
DB 1447 AATGCAATGCTTTGGCTGCGCTGACCGATTA 1479

RESULT 14
AX280370
LOCUS AX280370 2369 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 1 from Patent WO0177341.
ACCESSION AX280370
VERSION AX280370.1 GI:16607748
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
Smetts,G.M. and Sprengel,J.J.
Mammalian heparanase
Patent: WO 0177341-A 1.18-OCT-2001;
JANSSEN PHARMACEUTICA N.V. (BE)
FEATURES
source
1..2369
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 63.6%; Score 1132.2; DB 6; Length 2369;
Best Local Similarity 87.9%; Pred. No. 4.2e-306;
Matches 1321; Conservative 0; Mismatches 8; Indels 174; Gaps 1;

QY 277 CTGATTTCTTAAGTCCAGAGGCTTGTGTAACCTGCGCGGAGCTTGGCCGCTTT 336
DB 108 CTCTTCTCTGAGCTCCAGCGCTTGTGTAACCTGCGCGGAGCTTGGCCGCTTT 167
QY 337 CTGCGCTTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
DB 168 CTGCGCTTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227
QY 397 AAAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 456
DB 228 AAAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 287
QY 457 CGAAGTATGTTGCTTATGATTAACAGAAAGCTGCAAGATTGGCCAGACCTGATGT 516
DB 288 CGAAGTATGTTGCTTATGATTAACAGAAAGCTGCAAGATTGGCCAGACCTGATGT 347
QY 517 ATGCTGAGCTCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576
DB 348 ATGCTGAGCTCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407
QY 577 TTCTCAATTAATCAAGTATCTCATATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 636
DB 408 TTCTCAATTAATCAAGTATCTCATATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 440
QY 637 TTGCTGATGCTCTGAGAGCTCAACTGATATTTGCTTAATGACAGTGCCTGATATCC 696
DB 441 ----- 440
QY 697 AATACTCTGGAACAGTTCTAGTCCCTGAGTCTGTGAAGTACAGCGCCAGCAAAAG 756

```

[illegible]

Db	1434 TAA	1436
RESULT 15		
AF282885		
LOCUS	AF282885	1760 bp mRNA linear PRI 14-OCT-2000
DEFINITION	Homo sapiens heparanase-like protein HPA2a mRNA, complete cds.	
ACCESSION	AF282885	
VERSION	AF282885.1 GI:10801194	
KEYWORDS		
SOURCE		
ORGANISM	Homo sapiens (human)	
REFERENCE	Bukatyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1760)	
AUTHORS	McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R., Hitchcock, M., Patel, S., Barry, E., Stubberfield, C., Terret, J. and Page, M.	
TITLE	Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member	
JOURNAL	Biochem. Biophys. Res. Commun. 276 (3), 1170-1177 (2000)	
MEDLINE	20483645	
PUBMED	11027606	
REFERENCE	2 (bases 1 to 1760)	
AUTHORS	McKenzie, E.A., Tyson, K. and Stamps, A.	
TITLE	Direct Submission	
JOURNAL	Submitted (28-JUN-2000) Biology Group, Oxford Glycosciences, 10 The Quadrant, Abingdon Science Park, Abingdon OX14 3JS, England	
FEATURES		
SOURCE	Location/Qualifiers	
	1..1760	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/chromosome="10"	
	/map="10q23-q24"	
	61..1503	
	/note="splice variant"	
	/codon_start=1	
	/product="heparanase-like protein HPA2a"	
	/protein_id="AA033421.1"	
	/db_xref="GI:10801195"	
	/translation="MRYLCQAFPEAMPESNSRPPACIALPGALYLLALHLISQAGDRPRLPDAAGLKEKTLILDLVSTNPRATVENELSLDPSIHDGMLDLSKRLVLAAGSTVPAALNSGQKGYLQKSLDLPRIYRSALVGNPGRKRVIALDGFMAVASTVPAVTMHCYIDGRVYVMDPELTRELIDTSDQIRKQKRVNTHPEKTLISGVYTSAGGTRNLSRYAAGFLMNTLGMLANQIGIDVYIRHSFPBGRHNHLYDQNPFLPYWLSI LKRLGPRKVALAVHAGLQKRPFRGRVIRDLRLYACTNHNHNHYRGSITLTINILHRSRKKIKLAGTLEADKLHQVLLQAPQSGLSKRSVQLNQPFLVMDGTLPFLKPRRRLAAGRTLVIPVTMGFEVVKVNALACRYR"	
ORIGIN		
Query Match	61.6%; Score 1095.4; DB 9; Length 1760;	
Best Local Similarity	81.1%; Pred. No. 8.8e-236;	
Matches 1442; Conservative	0; Mismatches 1; Indels 336; Gaps 1;	
QY	1 ATGAGGGTGCTTTGTGCTTCCTCGTAAGCCATGCGCTTCAGCAATCCCGCCCCCGCG	60
DB	61 ATGAGGGGTGCTTTGTGCTTCCTCGTAAGCCATGCGCTTCAGCAATCCCGCCCCCGCG	120
QY	61 TGCGTAGCGCCGGGGGGCTCTACTTGGAGCTGTGTGCATCTCTCCCTTCCTCCAG	120
DB	121 TGCGTAGCGCCGGGGGGCTCTACTTGGAGCTGTGTGCATCTCTCCCTTCCTCCAG	180
QY	121 GCTGAGACAGAGAACCTTGGCTGCTGTAGACAGAGCTGAGAGTTTGAAGAAAGACCTG	180
DB	181 GCTGAGACAGAGAACCTTGGCTGCTGTAGACAGAGCTGAGAGTTTGAAGAAAGACCTG	240
QY	181 ATTGCACTTATGTAGAGACCAACAAGCCCAATGAGAGAGTCATGTGAATCTTCTCT	240
DB	241 ATTGCACTTATGTAGAGACCAACAAGCCCAATGAGAGAGTCATGTGAATCTTCTCT	300
QY	241 CTGAGCTGATCCGTCATATTCATGATGAGCTGAGCTGCATTTCTTAAGTCCAAAGCC	300

```

Db      301  CTGACGCTGGATCCCTCATCATATGATGCTGGATTCCTTAAGCTCCAAACGCG 360
Qy      301  TTGATGACCTGGCCCGGGGACCTTTCGCCGCTTTCGCGCTTCGCGGGGCAAAAGAC 360
Db      361  TTGGTGAACCTGGCCCGGGGACCTTTCGCCGCTTTCGCGCTTCGCGGGGCAAAAGAC 420
Qy      361  GACTTCTGCACTTTCGCAAACTGAGAACCCGCGGAAAGCCGCGGGGGCCCGGGCCG 420
Db      421  GACTTCTGCACTTTCGCAAACTGAGAACCCGCGGAAAGCCGCGGGGGCCCGGGCCG 480
Qy      421  GATTACTATCTCAAAAACATGAGATGACATTTTCGAGTGAATGTTGCTTAGATAA 480
Db      481  GATTACTATCTCAAAAACATGAGG----- 505
Qy      481  CAGAAAGGCTGCAGATGTCGCCAGACCCCTGATGTTATGCTGCTGCCAAAGGAGAG 540
Db      506  ----- 505
Qy      541  GAGGCTCAGATGATCTGCTCTCTTAAGAGACAAATTCOAATTAATTAAGTATCTC 600
Db      506  ----- 505
Qy      601  AATTTAAGCAGCAGGCTCTTAAGACAACTTATTAATTTGCTGATGCTGACCTCAC 660
Db      506  ----- 505
Qy      661  CTGATATTTTCTTAATGCACTGGGTGTAATCCCAATACTCTCGAAGAGTTCTAGT 720
Db      506  ----- 505
Qy      721  GCCCTAGCTGTTGAGTACAGCCGCAAAAAGTACAATTTCTTGGAATGCGGT 780
Db      506  ----- 505
Qy      781  AATGAGCCAAATTAATTAATGCAATGATGCGCGGGCAATTAATGCGAGCTGATGGA 840
Db      506  -ATGAGCCAAATTAATTAATGCAATGATGCGCGGGCAATTAATGCGAGCTGATGGA 840
Qy      841  AAGATTAATCACTGAGTGAAGAGCTGTTGCAAGCCATCCGAAATTTATTCAGAGCCAG 900
Db      555  AAGATTAATCACTGAGTGAAGAGCTGTTGCAAGCCATCCGAAATTTATTCAGAGCCAG 900
Qy      901  TTAATGAGCTTAATTAATGAGGCGCGCGAGAAAGAAATGTCATGCGCTCTTAATGATG 960
Db      625  TTAATGAGCTTAATTAATGAGGCGCGCGAGAAAGAAATGTCATGCGCTCTTAATGATG 960
Qy      961  ATGAAGTGGCAGAAAGTACAGTATGATGATGATGATGATGATGATGATGATGATG 1020
Db      685  ATGAAGTGGCAGAAAGTACAGTATGATGATGATGATGATGATGATGATGATGATG 1020
Qy      1021  CGGCTGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db      745  CGGCTGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Qy      1081  ATTAAGAAATTAATTAATGAGTATGATGATGATGATGATGATGATGATGATGATG 1140
Db      805  ATTAAGAAATTAATTAATGAGTATGATGATGATGATGATGATGATGATGATGATG 1140
Qy      1141  GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db      865  GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Qy      1201  TTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db      925  TTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy      1261  CACTATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db      985  CACTATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Qy      1321  GACTACTGCTCTCTCTCTCAAGGCGCTGATGCGGCCCAAGTCTTGCTGAT 1380

```

```

Db      1045  GACTACTGCTCTCTCTCTCAAGGCGCTGATGCGGCCCAAGTCTTGCTGAT 1104
Qy      1381  GTGGTGGGCTCCAGCGGAGGCAAGGCGGCTGCGGAGTATCCGGGACAACTAAGAT 1440
Db      1105  GTGGTGGGCTCCAGCGGAGGCAAGGCGGCTGCGGAGTATCCGGGACAACTAAGAT 1164
Qy      1441  TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500
Db      1165  TATGCTCACTGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1224
Qy      1501  ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db      1225  ATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284
Qy      1561  CTGGTTCACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db      1285  CTGGTTCACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1344
Qy      1621  CAATGATGCGGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db      1345  CAATGATGCGGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1404
Qy      1681  CGGCGCTTTCGCGCGCGGCGGACATTTGATGATGATGATGATGATGATGATGATG 1740
Db      1405  CGGCGCTTTCGCGCGCGGCGGACATTTGATGATGATGATGATGATGATGATGATG 1464
Qy      1741  GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
Db      1465  GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1503

```

Search completed: May 13, 2004, 14:18:53  
 Job time : 6978 secs